



Durham E-Theses

Developing concepts of musical style

Marshall, Nigel Andrew

How to cite:

Marshall, Nigel Andrew (2001) *Developing concepts of musical style*, Durham theses, Durham University.
Available at Durham E-Theses Online: <http://etheses.dur.ac.uk/3809/>

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full Durham E-Theses policy](#) for further details.

University of Durham

School of Education

Developing concepts of musical style

Nigel Andrew Marshall

2001

The copyright of this thesis rests with the author. No quotation from it should be published in any form, including Electronic and the Internet, without the author's prior written consent. All information derived from this thesis must be acknowledged appropriately.

This Thesis is presented for the degree of Doctor of Philosophy



- 7 JUN 2002

“...life is full of noise, death alone is silent: work
Noise, noise of man, and noise of beast. Noise
bought, sold or prohibited. Nothing essential
happens in the absence of noise”

Jacques Attali (1985)

For my dad,

Rev. Arthur Marshall

Who always knew I could do it.

Acknowledgements

I would especially like to thank the following people for their help and advice:

Firstly, Professor David Hargreaves of the Centre for International Research in Music Education, University of Surrey- Roehampton, who encouraged, helped and supervised the dissertation.

Also, Professor Barry Cooper, Department of Educational Studies at the University of Durham, who supervised the final stages of the dissertation, read and made helpful comments on the final draft.

Professor Carl Roberts. Department of Sociology and Statistics at Iowa State University, Ames. Iowa. U.S.A, and

Dr. Douglas Bonnett. Department of Psychology and Statistics at Iowa State University, Ames. Iowa. U.S.A who both helped with the Statistics, and

Dr. Stewart Thomas. Department of Psychology. Marshall University. Huntington. West Virginia. U.S.A. who provided further help and hardware for the statistics.

Dr. Kate Windridge (Castell) University of Leicester. Leicester. UK. for her comments and advice on her original research.

Mr. Roy Lidstone, Mr. Harry Elliott and Mr. Ron Harding helped by driving many miles *collecting and delivering papers, tapes and all kinds of other essentials.*

To my mother, Mrs. Irene Marshall for all her help marking papers, typing in data and checking final drafts.

Mrs. Wendy Bray, Headteacher and the Staff and Pupils of Fishbourne Cof E Primary School, West Sussex. U.K.

Mr. Bob Dabney and Mr. David Crockford and the Pupils of Teddington School. U.K.

Mrs. Leslie Reetz, Head of Music and the pupils of Fair Oaks Middle School. Fort Dodge. Iowa. USA.

Mr. John Mulett, Head of Music and the pupils of Fort Dodge Senior High School, Iowa. USA.

Mr. A. Lavery, Headteacher and the Staff and Pupils of Bishop Tufnell Junior School, Felpham. West Sussex. U.K.

Mr. A.J. Pratchett, Headteacher and the Staff and Pupils of Woodcote School, Surrey. U.K.

Mr. A. Simpson, Headteacher and the Staff and Pupils of Boxgrove Primary School.
West Sussex. U.K.

Mr.A.Lovett, Headteacher and the Staff and Pupils of The Vale School, Findon Valley.
Worthing. West Sussex. U.K.

Mrs. J. Wilson, Headteacher and the Staff and Pupils of Robin Hood School, Sutton.
Surrey. U.K

Mrs. V. Griffin, Head of Music, Robin Hood Junior School, Sutton. Surrey. U.K

Ms. Cheryl Sager, Music Teacher. Strafford School. Center Strafford, New Hampshire
and Hopkinton School, Concord. New Hampshire. USA.

Ms. Kate. Ingoldsby and Mr. John Green for their help with the testing.

Lastly, a second 'thank –you' to my friend Roy Lidstone, who spent many hours proof
reading several drafts and made many helpful suggestions

and to Lisa, who read and re read many of the sections, printed, e-mailed and saw it all
through to the end.

Abstract

This thesis explores the development of sensitivity to musical styles in children aged between 3 and 16 years old. The thesis is divided into four parts. The first part of the thesis explores the historical background of the developmental and the social psychology of music and reviews some pertinent previous literature. This section places the later studies in a theoretical context. The second part presents a review of the four previous studies, which have been carried out into style sensitivity, namely, Gardner (1973), Castell (1983), Tafuri et al. (1994) and Hargreaves and North (1999). This section of the thesis also includes a review of a number of methodological issues and gives a full description of the design of the test methodology.

Six experiments are carried out and reported in the thesis. Experiment one explores the effect of varying the lesson context in which the style sensitivity test is presented to the participants, whilst experiment two explores the effect of varying the test presenter on participants' test performance. The third experiment is a comparative study between participants in three different regions, one region within the UK and two regions within the USA. The participants in each of these three regions all experienced a number of variations in their music education programme, the status accorded to music within that music education programme and also considerable variation in participants' commercial musical diet. The experiment explores whether or not these variations affect the development of sensitivity to musical styles.

Part three of the thesis manipulates two further variables firstly, by varying the introduction to the style sensitivity and secondly, by manipulating the musical material used in the test. Experiment four presents two variations to the test introduction. In the first condition the introduction is changed through the offer of a reward and in the second condition, the introduction is changed in order to make the experiment competitive. The fifth experiment attempts to manipulate the musical material used in the style sensitivity test by increasing the level of stylistic divergence between the musical extracts. This is achieved in two ways. Firstly, a broad grained method of measuring various artefactual elements within the extracts is devised and incorporated into the selection of the musical test material. Secondly, the stylistic divergence between the musical extracts is increased by extending the chronological period between the musical eras used in the musical extracts.

Part four presents the findings of an exploratory study. Experiment six attempted to create a new test methodology which was appropriate for use on younger children. The new methodology was successfully piloted in order to further explore style sensitivity in participants aged between 3 and 5 years of age.

In the last section, a review of the six experiments is given and this is accompanied by a number of implications for further research. Finally, a model of the development of style sensitivity is presented. This model proposes that sensitivity to musical style is not a single unitary skill, the development of which is linear, and therefore any measurement of

musical style sensitivity should be seen as a multidimensional description of a combination of abilities, skills and knowledge. The model proposes that acquiring the ability to be sensitive to musical styles can best be seen as the progression from stylistic discrimination to stylistic competence.

Table of Contents

Abstract

Acknowledgements

Chapter 1: introduction

- 1.1 Historical background of music psychology
- 1.2 Studies of listening to music
 - 1.2.1 The use of musical stimuli
 - 1.2.2 Psycholinguistic influences in studies of listening
 - 1.2.3 Meaning in music
- 1.3 The present studies

Chapter 2: the developmental and social psychology of music

- 2.1 Theories of developmental psychology
- 2.2 Developmental theories of music
- 2.3 Two developmental theories in the arts: Koopman (1995) and Koroscik (1997)
- 2.4 The social psychological perspective
- 2.5 Historical background of social psychology
- 2.6 Developmental social psychology
- 2.7 The social psychology of music
- 2.8 Summary

Part one: development of style sensitivity

Chapter 3: literature review

- 3.1 Gardner's three systems model
- 3.2 Gardner (1973a)
- 3.3 Castell (1983)
- 3.4 Tafuri, Addessi, Luzzi and Baroni (1994, 1995, 1996)
- 3.5 Hargreaves and North (1999)
- 3.6 Summary

Chapter 4: methodological considerations

- 4.1.1 Research precedents
- 4.1.2 Experimenter rapport
- 4.1.3 Experimenter expectations

- 4.1.4 Pre test experiences and activities
- 4.1.5 Experimenter status and experience
- 4.1.6 Test anxiety
- 4.1.7 Experimenter distortion
- 4.1.8 Summary
- 4.2 The present studies: general methodology and procedure
- 4.2.1 Test design and modifications
- 4.2.2 Reliability and validity
- 4.2.3 Pilot studies
- 4.2.4 Musical materials in experiments one to four
- 4.2.5 Musical content in experiments one to four
- 4.2.6 Response sheet
- 4.2.7 Testing procedures
- 4.2.8 Marking procedure and analysis in experiments one to four
- 4.2.9 Method of analysis
- 4.2.10 Practice examples
- 4.2.11 Practice examples: music selection
- 4.2.12 Practice examples: marking procedure

Chapter 5: experiment one

The effects of lesson context

- 5.1 Background
- 5.1.1 *Listening environment*
- 5.1.2 *Prototypicality*
- 5.1.3 *Test environment*
- 5.1.4 *Musical environment and testing*
- 5.1.5 *Summary and background literature*
- 5.2 Design and procedure
- 5.2.1 *Schools*
- 5.2.2 *Participants*
- 5.3 Results
- 5.4 Discussion
- 5.5 Conclusions

Chapter 6: experiment two

The effects of variation of test presenter

- 6.1 Background
- 6.1.1 *Race*
- 6.1.2 *Experimenter gender*
- 6.1.3 *Experimenter age*
- 6.1.4 *Prestige effects*

- 6.1.5 Summary of background literature
- 6.2 Method
- 6.3 Design and procedure
 - 6.3.1 Schools
 - 6.3.2 Participants
- 6.4 Results
- 6.5 Discussion
- 6.6 Conclusions

Chapter 7: experiment three

National differences in musical style sensitivity: a comparative study in the UK and USA

- 7.1 Background
- 7.2 National and regional differences
- 7.3 Effects of training
- 7.4 The present study
- 7.5 Design and procedure
- 7.6 Schools
- 7.7 Participants
- 7.8 Results
- 7.9 Discussion
- 7.10 Conclusion

Part two: test instructions and style sensitivity

Chapter 8: experiment four

The effects of variation of test introduction on style sensitivity

- 8.1 Background
 - 8.1.1 Summary
- 8.2 Procedure
- 8.3 Schools
- 8.4 Participants
- 8.5 Results
- 8.6 Discussion
- 8.7 Conclusion
- 8.8 Summary

Part three: test content and style sensitivity

Chapter 9: experiment five

Variation in contrast between stimulus pairs: the effects of style and chronological period

- 9.1 Background**
 - 9.1.1 Previous empirical work**
 - 9.1.2 Theoretical perspectives on musical style and participant preferences**
 - 9.1.3 Analysis of musical style and participants preferences**
 - 9.1.4 Summary: The present study**
- 9.2 Selection of musical examples**
- 9.3 Measurement of stylistic divergence**
- 9.4 The assessment of musical examples: Condition Cc (contrast)**
- 9.5 The assessment of musical examples: Condition Ce (eras)**
- 9.6 Design and procedure**
- 9.7 Schools**
- 9.8 Participants**
- 9.9 Results**
- 9.10 Discussion**
- 9.11 Conclusions**
- 9.12 Summary**

Chapter 10: experiment six

Assessing musical style sensitivity in pre-schoolers: an exploratory study

- 10.1 Background**
- 10.2 Research on musicality in infancy**
- 10.3 Methodological issues**
- 10.4 Design and procedure**
- 10.5 Musical examples**
- 10.6 Participants**
- 10.7 Assessment procedure**
- 10.8 Results and discussion**
- 10.9 Conclusions**

Chapter 11: summary and conclusions

- 11.1 Introduction**
- 11.2 Summary of findings**
- 11.3 Implications for further research**

- 11.4 Discussion: Stylistic discrimination to stylistic competence - a proposed model of development**
- 11.5 Novice to expert**
- 11.6 Summary**

Appendices

Appendix 1 Musical samples: Experiments 1 to 4

Appendix 2 Musical samples: Experiment 5

Appendix 3 Musical samples: Experiment 6

Bibliography

Chapter One

Introduction

" Music has become more than a backdrop - it has become a blaring soundtrack for practically every event in our lives "
(Goodall, 1999, p.1)

Music has always been an essential, central and critical activity for all human beings. Evidence from bone flutes and cave paintings strongly suggests that primitive human beings made opportunities to indulge themselves in the human activity of making music. In spite of having very limited leisure time, music and the other human arts of painting and drawing, appear to have held an important position in all primitive societies. Authors have constantly argued over the reason why the earliest individuals involved themselves so much in the arts (Bronowski, 1973; Csikszentmihalyi and Robinson, 1990). Some writers (Read, 1955) have gone so far as to suggest that our ability to make and appreciate art and music "has remained a key to our survival" (p.32), whilst Attali (1985) reasons that without the distinctive way in which humans relate to and through sound, a truly human society may never have emerged. Whatever the reason, what we do know for a fact is that:

" experimental and hypothetico-deductive methods(are)
largely confined to Post Renaissance man: art antedates
recorded history"

(Gardner, 1994, p.24)

Our present day uses for music, the varieties of music and the meanings we attribute to individual musical activities are very different from the uses, styles and meanings used by our predecessors, even those of a hundred years ago, and therefore our modern day responses to music may also be very different. Different cultures and societies possess highly varied definitions about the constituent parts of music and what social purpose it should serve (Blacking, 1973; Gregory, 1997). Throughout the world many different musical systems exist. The rules of those systems and what the final musical product may mean or signify can be equally varied (Small, 1977). Similarly, musical roles may vary within and between each musical system and culture. Not all cultures have separate roles for composer and performer. Not all



musical cultures possess a concept of what it means to be 'musical' or 'non-musical'. Not all musical cultures divide the roles of performer and passive listener or audience member. In many folk traditions and also in some modern compositions (see Malcolm Williamson's work *Cessations* for example), there is no role for a passive listener and no audience exists – only a variety of performers. Likewise, the expectations of the responses from the listener, whether internal and private or external and visible may also be strongly regulated by the individual music system.

Even within one musical culture, individuals respond differently to individual pieces of musical stimuli; they may even respond differently to the same individual piece of music as performed by different musicians, and responses to the same piece can change over time, both in the long and short term. It has been demonstrated that individual reactions to musical stimuli may be influenced by many external factors such as social setting (Hargreaves and North, 1997; Konecni, 1982); typicality of the music to the situation (Martindale, Moore and West, 1988; Hargreaves and North, 1997; North and Hargreaves, 1996, 2000); the presence of others (Konecni, 1982); and physical factors such as age (Le Blanc, 1991); and emotional mood of the individual (Konecni, 1982). Factual information known about the music itself (Chapman and Williams, 1976) has been shown to affect the listeners' reaction to the music they hear. Konecni (1982) has shown that both pre-test and post-test activities such as insulting the test participants and allowing them to retaliate, can affect listener tolerances whilst North and Hargreaves (1996) have demonstrated that current activities whilst listening to music can affect listener tolerance, motor performance and even time estimation.

All these variables, both internal and external to the music, are potential subject matter for investigations within the field of music psychology. Hargreaves (1986) described the developmental psychology of music as a "newly emerging field of study" (p.1). Sloboda (1985) described a great gap that needed to be filled between the vast majority of psychological research on one side and the experience and insight of the active musician on the other (p. v). The ensuing years have seen a dramatic increase in the number of studies, the amount of research and the volume of literature which has appeared in every branch of this discipline. The original field of study has

now broadened well beyond the domain of the psychologist and the ramifications of present day findings are of interest to many other subject areas ranging from education to commerce. As Partch (1974) pointed out, everything in music, with the exception of the capacity of a body to vibrate and produce sound and the human ear that registers it, "is created by man or is implicit in human acts and is therefore subject to fiercest scrutiny" (p. 16).

This present study - 'scrutinises' - one of those 'human acts', namely the development of children's ability to perceive different musical styles. Writers such as Swanwick (1979) and Sloboda (1985) have suggested that musical acts can be classified into one of the following main categories: Composing, Improvising, Listening and Performing. In this present study we are concerned with the area of listening and this area of musical activity was chosen for a number of reasons.

1. According to the National Curriculum for Music (D.E.S.1992), all children should engage in, and have experience of improvising, composing, performing and listening to music. For the vast majority of children, their experience in all musical activities, with the exception of listening to music, will cease when they leave secondary education. Even those children who learn instruments and take an active role in the musical life of their school and their place of higher education, will cease to partake in any of these activities when they conclude their studies. However, almost everyone listens to music. Throughout their daily lives most human beings hear and / or listen to music. Knowledge about the listening process therefore can affect the greatest number of people.
2. Relative to other areas of music psychology there is still very little information about how stylistic sensitivity develops in children. In the 25 years following Gardner's initial study, only two further studies were published (Castell, 1983; Tafuri, 1994).
3. Studies of listening activities are valid areas for research. Gardner (1972a) argued that any study of children's listening processes would be valuable as it may further illuminate and identify deeper

cognitive processes and Blacking (1973) argued that the musical skill of listening was fundamental to music as a language and was too often ignored.

4. Recent studies (Hargreaves and North, 1997) have highlighted the numerous and interesting ways in which listening to music can be affected greatly by the social context. This research highlights an important new perspective from which the developing concept of style sensitivity should be viewed.

In this chapter we shall first look at the historical background of music psychology and examine some of the methodological issues and questions arising from this earlier work. This overview of the historical background allows us to identify three distinct areas of study that have developed within the psychology of music, namely cognition and perception; the developmental psychology of music and the social psychology of music.

We shall then review a number of previous studies that have explored various aspects of listening to music. These studies highlight a number of methodological issues surrounding the use of appropriate stimuli and problems that can be associated with the use of different types of musical stimuli. In the third section an overview of important developmental theories is given followed by a brief review of a number of theories that relate specifically to the development psychology of music. The growing importance of the social context of music is reviewed in the fourth section of this chapter and the final section reviews the structure of the study as a whole.

1.1 Historical background of music psychology

Early studies in the psychology of music, although not always attributed to that specific discipline, tended to focus on questions relating to the origins of music in humans. Dr. William Pole (1877) put forward the suggestion that music had evolved out of adult speech. This idea originated in an earlier suggestion made by Spencer (1857) and was published in Fraser's Magazine. Although Pole made a reference to

this earlier suggestion in one of his footnotes, he stopped short of any formal acknowledgement of Spencer. The suggestion made by Pole (and Spencer) was based on the observation that when adult speech became emotional, it tended to span a far greater tonal range and therefore became more like music. The fact that children were equally capable of making the same if not far greater variations in pitch during speech, and probably far more frequently, appeared to be of no consequence. Pole further suggested that words had gradually become separated from the sounds, which had then evolved into a sound language in its own right. Further arguments by Revesz (1953) and Blacking (1973) continued to suggest alternatives for the origins of music with some common factors emerging from the debate. There was some agreement that music and words had, in the past been far more closely linked. Music was seen as arising out of a strong, subjective emotional need to communicate with other individuals and arose prior to the later requirement of conveying objective information or exchanging ideas. More recent research on the use of Palaeolithic flints and the sound world of prehistoric man has again reopened the debate on the origins of music and its possible role in our evolution. For a full review see Cross (2001) and Tolbert (2001).

Other historical issues, apart from those concerning pure musicology, tended to centre on questions surrounding the philosophy of music and aesthetic problems such as the meaning and origins of music (Hanslick, 1986; Gurney, 1880). It is also clear that many early composers were well aware of the way in which 'perceptual tricks' could be employed in their compositions. Bach was certainly aware that the brain perceived certain patterns of sound according to primitive grouping mechanisms and frequently used these to full effect. For a full review and numerous examples see Sloboda (1985). Similarly, there were music educators who showed awareness of, and employed psychological phenomena in their music education schemes. (Carroll, 1910, 1914; Walker 1983, 1989).

Some earlier theories attempted a developmental account of music and also suggested the most effective form of education to achieve this. This approach is probably best expressed by Cecil Sharp (1907, 1912). Sharp suggested that folk songs were the natural, untainted and purest expression of musical development in a nation and

amongst its people. Sharp believed all forms of mental activity were due to the development and specialisation of qualities that were natural and inborn. Education, he claimed, could not create anything but could only develop those natural and instinctive faculties. He further argued that when innate aptitudes were given favourable conditions they would reveal themselves without the aid of formal education. He declared therefore that folk song was “the natural and untainted musical expression of a nation” (p.5). Sharp relied strongly on a theory borrowed from anthropology which stated that every individual in the course of his development had to pass through, and in the same order, all the stages that the human race had ‘painfully and laboriously traversed through’ over the centuries. Sharp therefore stated that music education for young children should contain only folk song, followed by the national songs and then last of all they should be taught part songs, notation and the science of music. Sharp completed his argument for music education with the warning that:

“ To invert this order, or to proceed prematurely to the study of part songs or of notation, is to commit a grave educational error” (p.13)

Cox (1990) has argued that through a figure such as Sharp, music should have become a radical educational force. Instead, for a number of reasons it became a regressive influence for three main reasons. Firstly, there was an over reliance on evolutionary ideas borrowed from anthropology with nothing based on studies in psychology. Secondly, there was a lack of sound educational goals stemming from Sharp’s view of childhood as being a primitive evolutionary state which would evolve more according to the rules of evolution than education. Lastly, Sharp has a tendency to regard folk music, and therefore music education, as a means of social control. Sharp was autocratic and it was he who dictated what constituted a ‘folk song’ and his overwhelming desire was not to promote musical skills and abilities, but to maintain the art of the rural communities in the new industrial society.

“ Perhaps the folk had become a middle class construct, expressing a naïve romanticism at best, a form of social control at worst.”

(Cox, 1993. p.159)

The major legacies left by Sharp are that he raised the profile of music as a curriculum area, although possibly for the wrong reasons. The board of education accepted his view that music was a social force and had beneficial effects upon the character. Secondly, along with Walter Carroll and others, he was a major force behind music education becoming available for the masses. Lastly, the composer Gustav Holst (Karpeles, 1973) claimed that Sharp would not be remembered for the music he had made himself but for the music he had allowed others to make. A number of further criticisms of Sharp as a force in music education have been presented elsewhere notably by Harker (1985); Middleton (1990) and Szczelkun (1993).

Although a number of individual studies explored some isolated music / psychological issues (Heinlein, 1928), probably the first major empirical basis for investigating the psychological aspects of music began with the work of Seashore (1938). Seashore adopted an essentially psychometric approach and was interested principally in measuring musical ability, aptitude and the pure measurement of the dimensions and capacities of the human ear. Although his book was entitled *Psychology of Music*, the contents were mainly related to the physics of sound, acoustics, individual instruments and even the physiology of the performer. Seashore developed a battery of tests (*Measures of Musical Ability*, 1919) that were designed to measure the capacity of the human ear to recognize the changes in pitch, rhythm and timbre. He believed that the sensory capacities of the individual could be identified and measured and the extent to which these capacities were either present or absent formed the foundations of the measurement of that individual's musical aptitude.

Seashore tests were later criticized from a number of sources (Davies, 1978) with the main criticisms falling into three broad categories. First, there was criticism of the way in which the test materials used electronically generated tones and not examples of actual music. The test examples therefore lacked those human elements which are normally present in more natural musical examples. Elements such as variation in tempo and lack of phrasing which are present in more natural examples may provide additional data, which the listener may use in fully interpreting and understanding music. Secondly, it was argued that the tests of musical ability did not satisfy criteria

of reliability and validity, they were therefore seen as being methodologically flawed (Drake, 1939; Teplov, 1966). Thirdly, the exact relevance and relationship between measures of raw sensory capacities and musical behaviour was questioned and as a result of this questioning, much broader definitions of musical ability were introduced (Revesz, 1953). A full summary of this debate is given in Lundin, 1967. Revesz (1953) and Lundin (1967) argued that participants who were musically talented and competent were more likely to achieve poor scores on any test material that incorporated non-musical or synthetic sounds. This is because these artificial sounds did not contain many of the human features that most musicians / musical individuals, are highly attuned to, such as phrasing and rubato.

One major benefit that resulted from Seashore's work was that his methodology, reported work and published battery of musical ability tests did appear to act as a catalyst for the debate on both the nature of musical ability and the methodological issues surrounding any test or study of musical effects; a debate which still continues today (Persson and Robson, 1995). Seashore had himself made references to the limitations of his work. He pointed out "the musical mind does not consist of its dissected parts, but is an integrated personality" (p. 1). Seashore highlighted the possible effects caused by the differences between human and musical aspects and the pure physical aspects of any sound. He noted the difference between the frequency of a note, which he was able to measure accurately in a laboratory; and the human ability to recognise pitch, which is subjective and dependent on a number of cultural, experiential and possibly physical variables.

Seashore was attempting to measure the purest aspects of the ability of the human ear. He held the belief that there existed hereditary 'fundamental abilities'. Although he acknowledged that environmental factors could and would influence musical ability, he still argued that it was possible to isolate and measure these hereditary fundamental 'raw' abilities. He therefore used pure tones in his experiments because, he argued, these would measure this 'raw' musical ability of the individual. The fact that the tones were unmusical removed any possible effects caused by previous training, home background or previous musical experience. The measurement of this 'raw' ability, untainted by any previous experience would, Seashore argued, be a clear and accurate

indication of those individuals who would gain most benefit from music education programmes.

Seashore's psychometric work remained prominent during the years from 1950 to about 1969 and it gave rise to a number of musical ability tests such as the *Watkins - Farnum Performance Ability Scale* (1954); the *Colwell Music Achievement Tests* (1969-1970) and the *Tests of Musical Ability and Appreciation* (Wing, 1968). These tests followed Seashore's original pattern by presenting a battery of tests, each of which examined only one specific component of musical ability. Even today there are a small number of authors who continue with this approach. *Gordon's Music Aptitude Profile* (Gordon, 1979) is based on the notion that measuring raw levels of musical ability can produce valid measurements of musical ability upon which educational decisions may be made.

The debate generated by the work of psychometricians such as Seashore and Wing also influenced the development of a number of more holistic approaches. For example, Bullock (1973) reviewed tests that aimed to assess a variety of perceiver attributes which are more general, and hence less easy to define than specific abilities such as pitch or rhythm perception or discrimination. Bullock suggested that a large number of influences e.g. mood and interest, may contribute to what he termed the total "musico-aesthetic attitude"

Further and more recent studies acknowledged the influences of other disciplines such as ethnomusicology and sociology and began to bring about an evolution in the psychology of music into a number of specific research areas. Hargreaves (1986) details clearly the rapid increase in both the volume of research and the amount of published literature which appeared during the early years of the 1980s, in fact, "a new textbook on some aspect of music psychology has been published almost every successive year" (p. 3). Since the 1980s, work on the psychology of music would appear to have grown into three main areas of study. The first area concentrates on research in the cognitive psychology of music (Sloboda, 1985) whilst the second area is concerned with the developmental psychology of music (Hargreaves, 1986). Within

the last few years a third and new area has begun to emerge which concentrates on the social psychology of music (Hargreaves and North, 1997).

The first of these areas explores the cognitive elements of music such as musical perception and memory (Sloboda, 1985). In the Preface to his book, Sloboda spoke of the gap existing between the study of music psychology and academic musicians, neither group seemingly willing to cross the "interdisciplinary barrier". *The Musical Mind* was an attempt to review the already existing body of literature and propose a number of research questions which may begin to bridge this gap. Sloboda argued that individuals participated in musical activity because of the levels of arousal and the level of emotion, which humans experienced from music. The link with cognitive psychology, he argued, came through two pathways; firstly, by virtue of the fact that most of our responses are learned and secondly, because the cognitive component of our emotional response could not be adequately explained through the idea of 'conditioning' (p.2). Sloboda used this perspective from which to explore the cognitive elements of performance, composition, improvisation and listening.

The second area explores the developmental psychology of music (Hargreaves, 1986). In *The Developmental Psychology of Music*, Hargreaves argued that music psychology had previously consisted of a number of loosely related topics that lacked a suitable conceptual model or theoretical framework. He further argued that developmental psychology had previously been synonymous with child psychology with many studies being simply descriptive rather than explanatory and lacking any theoretical framework. He noted that there were:

“.. virtually no coherent psychological theories of the specific developmental processes underlying children's musical perception, cognition or performance” (p.3)

He commented also on the gap, which, he noted, existed between the theoretical aspects of music psychology and pedagogical studies surrounding the best teaching methods for music. Hargreaves further argued that there had been a considerable lack of interest in the development of musical skills throughout childhood. The majority of studies exploring the development of musical skills had all tended to focus on the differences which existed between adult participants with differing levels of musical

skills rather than on psychological changes occurring in young developing children. There were, however, a number of developments and changes within the discipline of psychology, which substantially influenced the developmental psychology of music, assisting it to become the thriving and clearly defined paradigm which it now has become. Firstly, there has been a dramatic increase in technological developments and abilities now available to researchers. More accessible and powerful personal computers and more innovative software have enabled more accurate and sophisticated data to be collected and processed. Secondly, there was an increase in the understanding of developmental research strategies (Schaie, 1965), and thirdly, there has been a large increase in imaginative research strategies, often involving the technological innovations already referred to, such as computers linked to video cameras and audio equipment. These advances have enabled researchers to gather and analyse large amounts of multidimensional data on infants and children in ways that had previously not been possible.

The third and most recent strand to emerge is the social psychology of music. This approach acknowledges that musical events do not take place within a social vacuum (Farnsworth, 1954,1969; Hargreaves and North 1997). The growth of this area of study has been due to a number of factors. Firstly, there is the ever-increasing use of music in our present day society. Dramatic advances in technology, the growth of the music industry and the affordability of personal audio technology has provided individuals with the opportunity to hear an ever-increasing number of types and styles of music in an ever increasing number of situations. Secondly, Hargreaves and North (1997) suggest that developments within the disciplines of psychology and social sciences have increasingly begun to accept that social and cultural context in which an event takes place, can be a vital and informative factor. Likewise, developmental psychologists have begun to accept the importance of the social context within the process of human development. Just as any musical activity takes place within a social context, so human beings develop within a social context. The links between the two cannot, it is argued, be isolated from each other.

Further issues surrounding each of these three specific areas within the psychology of music are explored more fully in subsequent sections.

1.2 Studies of listening to music

In this section we review the major areas of research on listening to music. The studies briefly reported in this section raise a number of methodological issues surrounding the problems and benefits of using different forms of musical stimuli; natural versus experimental approaches, and the use of short versus prolonged excerpts. A number of methodological issues, particularly those concerned with short musical extracts and the breaking down of music into smaller component parts, were informed by research into language perception and cognition. We therefore include a brief section on how the study of linguistics has informed and shaped some of the research on listening to music.

The value of listening to music has been emphasised many times by musicians, music educators, musicologists and psychologists. More people take part in musical activities in the role of a listener than any other, in fact, it is hard to find any individual who engages with the world, who does not take part in a significant amount of musical listening. Yet as Aiello (1994) points out, there is a significant lack of research on prolonged listening activities when compared to the amount of research on other musical activities such as composing, improvising and performing.

".. music psychologists have been reluctant to investigate how listeners perceive actual pieces of the repertoire. Therefore, little is known about the processes that take place while listening to a composition" (p.274)

Blacking (1973) argues strongly for the status of the role of listener in musical traditions and acknowledges the level of skills required to fulfil this role

" And yet the very existence of a professional performer
..... depends on listeners, who in one important respect
must be no less musically proficient than he is" (p. 9)

Blacking emphasises that creative listening is ignored too often in discussions of musical skills and abilities. Blacking argues that the study of listening skills should not only inform us about mental processes but should elevate the activity of listening to the same status as that accorded to composing and performing, as is the case in other cultures. Gardner (1972a) is also keen to stress the importance of researching

the activity of listening. He suggests that research on listening will not only inform us as to how people perceive music and develop musical skills, but that knowledge gained about the process of listening to music may well illuminate the way in which other deeper cognitive processes function.

Listening is a valid and valuable musical activity that is entered into by more people than any other musical activity. Any research that adds to our knowledge of the processes involved in listening, the effects that listening to music has on individuals and how education can best serve the requirements of the creative listener will be valuable to psychologists, musicians and educators.

1.2.1 The use of musical stimuli

Listening studies that incorporate different forms of stimuli have created a substantial amount of debate with the strengths and weaknesses of each form being strongly advocated by its' proponents. However, there are three general issues that may affect all studies of listening to music. The initial difficulty in studies of listening relates to the fact that listening, as a musical activity, often does not involve the production of any tangible or observable end product in the same way that composition or performance does. It is therefore difficult to tap into and explore the moment-by-moment activities taking place in the listener as they hear musical stimuli. Secondly, observations of minor physical responses that do occur in response to musical stimuli, for example dancing or clapping, provide only limited information about the total listening process. Such responses would be identical to any number of musical stimuli incorporating similar elements such as rhythm and therefore tell us little about the processing of artefactual elements within that specific piece of music. Lastly, studies that depend on the subject talking about the listening experience after the stimulus has finished are dependent on memory recall. As Sloboda (1985) suggests, despite being fully involved with a piece of music at the time of the actual performance, it is frequently impossible to recall major themes and features from it after the piece has finished. Similarly, studies involving verbal responses from participants rely on the use of language. The listening response is therefore to some extent, mediated by language capacity.

Many studies of listening to music have often used short stimuli involving either one or two notes (Seashore, 1938), or one or two short contrasting passages often containing no more than twelve notes (Deutsch, 1982; Dowling, 1973). There are three main methodological issues with this. Firstly, it has been argued that using short stimuli of computer generated or specifically composed short musical sequences is beneficial. Using this type of stimuli can give psychologists some control over variables within the sequence. Using this approach, studies have concentrated on the components of music such as melodic contour, harmonic structure and rhythm. (Seashore, 1938; Bamberger, 1979; Pflederer Zimmerman, 1967). However Aiello (1994) has argued that this reductionist approach has reached a natural ending and broader perspectives must now be taken in order to increase our knowledge and understanding of how music affects us and how we process music as a symbol system. Aiello suggests that the process of listening to music in the 'real world' is very different from the mental activity employed within studies that employ only short stimuli.

Secondly, music is more than a sum of its constituent parts. Musical segments are dependent on musical context and our processes and reactions will vary according to variations in that context. For example, if a subject experiences a short segment of melody played on a single instrument within a listening experiment, that melody will be perceived, processed and reacted to in a number of ways. That identical melody may cause very different reactions when it appears within a prolonged piece of music where it is recurrent, embellished, developed and played in a number of ways by various instruments. Often the full meaning of a melody can only be seen in terms of the way it is developed and presented throughout the overall piece. Studies involving short sequences for stimuli take no account of this.

Third, music is something that is experienced over a temporal continuum. It could be argued that prolonged musical pieces simply contain a large number of smaller segments linked together. However, as Salzer (1952) points out, the understanding of a musical piece often depends on the listener connecting a number of musically similar segments, which occur at some distance in temporal terms. In using any

segment of a much longer piece of music, certain vital relationships within the music, which are evident when played as a whole, are missing when the listener simply hears a brief excerpt. This can certainly apply in the case of certain styles such as Classical where the actual form of the piece is an important feature of the style and yet is not identifiable within the segment selected.

The use of short segments of musical stimuli can be of use in informing us as to how individuals perceive and process some forms of musical information. However, a degree of caution is required when we are considering the extent to which these results do inform us about other listening activities. Furthermore, simply using prolonged musical stimuli may not address the concerns noted here.

A number of researchers have attempted to address these issues through studies that incorporate prolonged and natural musical examples (Pollard-Gott, 1983; Clarke and Krumhansl, 1990; Deliege and El Hamadi, 1990 and Deliege, 1989). Pollard-Gott investigated whether listeners' perceptions of themes within a piece changed with subsequent hearings. Participants were required to take note of their responses during sessions of listening to music, and it was found that the perception of thematic material did change with multiple hearings and also that musically trained participants appeared to perform better than musically naive participants.

Clarke and Krumhansl (1990) have investigated how musically experienced participants segment music, focusing on how they identify boundaries, locate excerpts within the overall piece and estimate their duration. Deliege and El Hamadi (1990) have considered the perception of musical form. Their data revealed two organising principles were used by listeners; that of sameness and that of difference. The principle of sameness helped them constitute groups out of what they heard, while the principle of difference helped them to differentiate what they listened to. Both musicians and non-musicians showed important similarities in their perception of the composition.

Aiello, Tanaka and Winbourne (1990) carried out a number of studies that explored musicians' listening processes. Their study explored what musical features the listener selected to hear as a prolonged and complex musical piece developed over time. Participants heard the piece twice through. During the first hearing, participants were required to listen attentively whilst during the second hearing they were asked to note down any musical element they heard. Participants were music graduates and all were familiar with the piece. The results suggested very different listening patterns amongst the participants with some concentrating on very specific and segmented details whilst others adopted a more general and holistic approach. Results did not appear to be significant with respect to participants main instrument. All participants did, regardless of their overall approach, comment on the more salient items of the piece such as set chord sequences and tonality. Follow up studies made no further claims. Participants appeared to respond less to subsequent hearings, the authors suggesting that boredom with the task was a possible reason. Also it was noted that a greater quantity of responses were made to Mozart than other composers.

Aiello (1994) is a strong advocate of the need to use prolonged excerpts of 'real' music in studies of listening. She argues that studies involving short or artificially produced stimuli dramatically reduce the number of listener choices. This occurs in two ways. Firstly the amount of musical material is reduced and therefore the number of choices the listener can make in terms of what aspect of the music to focus on is also reduced. Secondly, the number of possible responses the listener can make is also reduced. For example, highly emotional responses are hardly likely to be possible. Aiello therefore argues for the use of prolonged and natural musical examples to be used in studies of listening because this restores the listener choice and further examination of these choices may provide us with a more accurate picture of how individuals process listening to music.

Budd (1985) suggests there may be a difference between the activities of *listening* and *hearing*. Hearing music takes place when listening to music is not the sole and primary aim of the experience. So, for Budd when music is danced to, marched to or the listener is "engaged in some other activity that occupies all or part of his attention" (p.16), that individual is hearing the music but not listening to it. According to Budd,

the individual who is hearing can be affected by some elements of the music, for example, rhythm, when marching or beat when dancing. However, the processes at work may well be very different from those processes being employed by the listener. As Budd points out, the listener is more informed of his role and:

" he listens to what he hears with expectations he would not otherwise have" (p.16)

Budd uses vivid and physical examples (dancing and marching) as representative of other activities, which the listener / hearer may be involved in. However, even within conditions that could be described as being 'listener orientated', individuals can often become involved in other 'extra' activities. Within the concert hall for example, such other activities as watching musicians, moving or tapping feet or sub-conscious interactions with other audience members may occur. These activities may in turn create a slight yet significant shift along the continuum from 'listening' towards 'hearing'.

Janet Mills (1998) has also discussed the difference between 'listening' and hearing' and carries out a further review of the 'new vocabulary' which has developed in accordance with the new concepts required by music educators. Mills (1998) tends to agree with Budd (1985) in that 'listening' involves some degree of intent and concentration whereas 'hearing' is essentially, a passive act. Therefore, from a formalist perspective, 'hearing music' could be seen as a problematic concept, because there is no active processing and understanding of the forms and structures of the piece in question. Cook (1990), taking a mainly formalist perspective, suggests that no musical interaction takes place when an individual 'hears' a piece of music. However, this assumption does not take account of the fact that music which is 'heard' does appear to have the power to influence the mood / actions of the 'hearer'. North, Hargreaves and McKendrick, (1997) found that supermarket shoppers who 'heard' certain musical styles (French accordion and German bierkeller music) were influenced in their choice of wine purchase. The results from this study suggested that some form of 'low level interaction' did take place, even within the mind of the 'hearer'. Hargreaves et al. (1997) have therefore proposed the idea that there are possibly 'levels of engagement' which exist along a continuum. 'Listening' can be said to involve high levels engagement and therefore high levels of interaction with

the music, whilst 'hearing' would involve low levels of engagement and thus low levels of interaction with the music. However, the research by North, Hargreaves and McKendrick suggest that some form of interaction can take place, even with low levels of engagement.

This view therefore suggests that the difference between 'hearing' and 'listening' is perhaps not distinct and precise, but varies by degree of engagement. The level of engagement, and therefore the interaction with the music, may even vary and move along the engagement continuum frequently throughout one individual piece of music as the individual interacts more or less with different sections of the same piece.

With specific reference to the study reported in this thesis, the idea of a continuum of interaction appears to be far more appropriate than the idea of two distinct aural activities. Accordingly, if the findings of Castell (1983) are correct, then participants will be operating far more towards the 'hearing' end of the continuum with the classical pairs, and far more towards the 'listening' end, with the popular pairs.

One study by Deliege and El Hamadi (1990) employs prolonged musical stimuli and incorporates the use of a timing mechanism linked to computer. In this study the subject was asked to record the moment when they perceived "segmentation" occurring within a piece of music. The subject recorded the exact moment by depressing a key. The computer recorded this. However, this form of experiment must be subject to the same criticism made by Aiello in that the listener is again directed towards those aspects of the music the experiment is designed to explore. In addition participants are also required to engage in some form of computer activity. The subject is therefore operating, as Budd says, with expectations which they otherwise may not have had. Pollard-Gott (1983) gave participants instructions including what type of musical element to focus on. Participants also had to respond according to an 11-point scale. Again, total listener choice has been removed and listener responses may be brought into play as a direct result of either the verbal instructions for the experiment or the experimental activity itself.

Therefore, the use of prolonged musical stimulus can also be the subject of methodological criticism. Experiments utilising prolonged or whole pieces of music involve the subject either reporting as the music is heard, which can be said to a different listening activity than the same subject sitting alone; or the subject is required to respond after the event, which may be subject to memory and recall difficulties as detailed earlier.

In both cases participants are also directed towards specific aspects of the music. They cannot therefore be said to have total listener choice.

A review of the literature suggests that there are difficulties inherent in studies of listening to music whether they incorporate short or prolonged stimulus. As Aiello (1994) points out in many studies:

“the participants listen to ‘real’ music, but their responses may or may not duplicate their actual experience when listening to music in a non experimental situation” (p.280)

In conclusion, studies involving listening to music must obviously involve the use of musical material presented in either short, artificial segments or in prolonged excerpts selected from full scale musical works. Each form of stimulus is subject to its own form of methodological criticism. Yet studies utilising each form of stimulus do present interesting and unique ways of viewing the way in which humans perceive, process and respond to music. Any weakness in a particular ideology may therefore lie not within the stimulus itself, but within any attempt to generalise the findings beyond their original context.

The study of linguistics has been an important influence on the study of music listening. Psycholinguistics has informed the design of research studies; has raised important questions about methodological issues in studies of listening to music and lastly, has shaped how some writers view the development of musical abilities. In the next section we explore the influence of psycholinguistics by looking at how some writers have explored the interesting similarities between language use, language development and music.

1.2.2 Psycholinguistic influences in studies of listening

Studies in the psychology of music have been influenced by developments in psycholinguistics. Theories surrounding the genesis of music (Pole, 1877; Sachs, 1943) alongside the similarities between language and music have prompted some researchers to investigate the perception of music by using similar experimental designs to those used to investigate language (Clarke, 1989; Lerdahl and Jackendoff, 1983). Sloboda (1985) reviewed a number of areas in which there appeared to be a degree of similarity between language and music. Both language and music are universal and specific to human beings. Both language and music are capable of generating an unlimited number of sequences and infants are able to learn the rules and to understand both through natural exposure. Both language and music include the use of voice and hearing and therefore many of the neural mechanisms for analysing input and producing the output may be shared; the most universal expression of all musical forms is the song. Aiello (1994) develops this musical – linguistic link by detailing how a number of languages use pitch contours to change the meaning of words.

Sloboda also points out that many cultures have developed often quite sophisticated methods of recording and notating both language and music and, although the forms taken by music and language vary across very different cultures, a number of universal parameters exist which shape, constrain and define music e.g. the arrangement of tones into culturally accepted scales. Sloboda has also examined the relationship between the phonology, syntax and semantics of music and language. He points out that there is some evidence to suggest that the same perceptual mechanisms are possibly present in both speech and in musical understanding. Experiments have shown that participants exposed to short sequences of language or music over which an audible click has been superimposed show a tendency to 'move' the click to the boundaries between structural groupings (Fodor and Bever, 1965; Gregory, 1978). Similarly, the tendency shown by participants to pause longest after fixations which complete major grammatical units in reading literature, (Green and Mitchell, 1978) has been shown to also exist in music. (Sloboda, 1974, 1977)

Psycholinguists distinguish between three components of language, namely: phonetics - that is the sounds of a language, syntax - that is the way in which the sounds are placed together, and semantics - the meaning of individual words. In order to understand language fully, the listener must have the ability to sufficiently process all three components. As has been stated, there are a number of similarities between language and music that have been observed. These observations have shaped a number of experimental techniques and these techniques have in turn produced a number of further similarities. Research into the phonetics of language and music has explored the way in which sounds are categorised. Perception takes place partly through categorizing sounds. The ability to categorize sounds has been found to be greater in trained musicians than in non-musicians in the same way as native speakers of a language categorise speech sounds more accurately than non-speakers (Repp, 1984). Locke and Keller, (1973) demonstrated this through asking musicians and non-musicians to categorise major and minor triads. Non-musicians were found to require a far greater frequency difference between the two chords before discrimination could take place.

Another similarity between music and language was established by De Witt and Samuel (1990). They demonstrated the phenomena of phonemic restoration; that is the ability to use previous knowledge to supply missing information. This ability has been found to exist in both language and music. These authors demonstrated that when participants hear a word with one small section of sound distorted, they are able to fill in the missing information from their previous knowledge of the word. Similarly in music, participants are able to draw on their previous knowledge of musical structures in order to fill in individual sounds that are missing or distorted. Participants believe they have heard a complete sequence when in fact they have themselves replaced some of the notes. Sloboda (1985) has also reported on literature which supports this finding. One experiment was carried out on both professional and amateur musicians. Each group was asked to play a selection of musical excerpts. In some of the excerpts, certain single notes were deliberately omitted from certain passages, for example one note was omitted from an otherwise complete scale passage, or a single note was omitted from an otherwise complete glissando. The experiment suggested that the more professional players were more likely to read and

play the passages incorrectly by including the missing note. That is, they inserted the missing note and therefore played what was more usual and not what was written.

Research into syntax, the order and way in which words are placed together, has also informed some musical studies. Results suggest that as in language, the non-musically trained listener still acquires the culturally agreed set of rules that govern music within a set style. Deutsch (1982) showed that participants were far more able to recall 12 note sequences when they were placed in some form of musical structure that closely followed the rules of music most familiar to the participant's experience. The rules of language, or grammars, are strategies for generating and understanding sentences and a number of researchers have investigated if musical grammars exist (Lindblom and Sumberg, 1970; Lerdahl and Jackendoff, 1983).

Psycholinguistics has informed and shaped a number of studies in the psychology of music. These studies have produced interesting and informative results but psycholinguistic paradigms can only produce a limited analysis of the total picture of the psychology of music. The production of and understanding of music depends on much more than simply following and understanding a grammar; indeed it can be argued that the most successful of creators / composers / artists, are those who move beyond or disregard rules and grammars.

In contrast to the studies reviewed briefly above, Aiello (1994) comments on the differences between language and music. She suggests that music is far more flexible and ambiguous than language and concludes that whilst psycholinguistics has assisted early studies in music psychology, advancing research must now take note of the unique aspects of music if a more accurate picture of how we process data when listening to music is to be produced.

Cook (1990) has produced a number of interesting questions around the relationship between language and music and also the communication of musical meaning to the listener. Cook (1994) challenged a number of assumptions made by previous studies in the perception of music and further challenges the usefulness of psycholinguistic

influences, and directs his main argument specifically towards the work of Wolpert (1990), although the criticism can be generalised to many others. Briefly, Wolpert presented musicians and non-musicians with three musical excerpts. First, tune A played on instrument X; then the same tune A played on a different instrument Y. Lastly, they heard a different tune B but played on the first instrument X. Wolpert then asked participants which two of the three tunes were most alike. Musicians all grouped their decisions around the same tune on a different instrument whilst non-musicians chose the same instrument playing a different tune. Wolpert then carried out a similar study with the same tune played on two different instruments for two of the examples. The third example consisted of the original tune and the original instrument but the accompaniment was played a fifth below the tune. That is, the tune and the accompaniment were in different keys but on the original instrument. Non-musicians again selected the two excerpts played on the same instrument as being most alike, in spite of the fact that the second example had the accompaniment played in a different key. Identical instruments outweighed any differences in accompaniment. On the other hand, musicians opted for the different instruments but with the same tune and musical accompaniment. The results led her to conclude that musicians and non-musicians do not listen to music following the same rules. She suggested that:

“ Musicians consistently chose melody and correct harmony accompaniment over instrumentation when matching an excerpt to a model, but non musicians did not do so” (p.103)

That is, musicians focus on what makes most musical sense out of the whole musical excerpt, whereas non-musicians focus on perhaps more superficial and dominant issues such as instrumentation regardless of other perhaps more subtle elements which may not make musical sense. Wolpert argued that her results showed “ a profound overestimation of what it is that most listeners hear” (p.103) and further more, music educators may need to teach a new awareness of harmony and melody. In other words, only musicians may be hearing what the composer intended.

Cook (1994) argues that musicians trained in the Western tradition are involved in a language game. “This ‘game’ tends to be organised around ‘things’ to play, with a ‘thing’ to play on’ ...” (p.69). Simply, musicians and non-musicians detect and are

equally aware of the two differences in Wolpert's experiment, however, they choose to answer musical questions in different ways with different priorities. Musicians answer musical questions in terms of the language game of Western music. Non-musically trained individuals do not have such a language. The training individuals receive in order to become musicians also trains them to answer musical questions. Wolpert extends her claim further. Her argument suggests that as musicians and non-musicians hear things differently, it therefore follows that:

“There is an immense gulf between the creative intent of the composer or improviser and the perception of the audience” (p.104)

That is, the true and full meaning of music is denied to all non-musicians. For Cook, this raises the additional question of how far the intent of the composer is realised in the perception of the listener. Essentially, Wolpert's argument is that an individual is required to be trained as a musician before they are able fully to perceive the intent of the creator of a piece of music.

Cook (1987) attempted to explore this claim by investigating the perception of one element of a musical piece over increasing periods of time. Using both musicians and non-musicians he carried out a series of studies in which he investigated ‘tonal closure’: that is, when musical pieces establish a home key (tonic). Excursions to other keys represent increases in arousal and or tension. Tension is resolved by a return to the stability of the home key; that is tonal closure. Cook's experiment explored listener preference for the same piece of music where the resolution of the tonal closure was delayed over an increasing amount of time. His results suggested that even in trained musicians, there was an optimum time limit in which tonal closure could be delayed. The fallibility of memory and the ever-increasing amount of data to be processed prevent the recognition of the tonal closure. That is, in perceptual terms, an extended composition cannot have the tonal unity of a single phrase. This suggests that large-scale structures cannot be understood in the same way as small-scale structures, at least in the way proposed by Schenker. (See Sloboda, 1985)

Cook adopts a criticism made by Harris (1980) of psycholinguistics as a whole and applies it to listening to music. Cook challenges the notion that a musical score is a representation of the sound one hears when listening to the score being ‘performed’. He argues that notes are not necessarily ‘in’ the music but are constructed out of it. He

suggests the example of a singer who sings a rapid scale. Essentially, there are no 'notes' as they appear on the score. Physically, the singer produces something closer to a glissando, sliding from one note to another. However, the individual 'notes' are placed 'in' by the listener who hears according to cultural rules.

Cook (1994) lastly challenges the idea that the cognitive processes involved in composing/ improvising and listening are essentially the same. That is, in order for the music to communicate fully it's meaning, the composer and listener must follow the same rules and processes. According to this view, the musical process of creation, transmission, reception and cognition becomes more like a secret service operating in code. The composer 'scrambles' the original code, transmits it through performance and the musical listener, having a copy of the identical code 'unscrambles' and therefore deciphers the 'message' at some other point in time and place.

Cook (1994) suggests that this notion whereby the individual receives the *message* by decoding that which the composer has encoded, is based on a number of disputable assumptions. The first assumption is that the listener chooses to listen in an analytical way, searching for musical grammars. The second assumption is that there is, or should be a musical grammar and cognitive process that is common to creator and listener and lastly, the assumption that even within the work of one composer, a central musical grammar actually exists. Musical theory codifies and institutionalises musical style. Musical theory can explain music but not in the same way that formal grammars explain language.

To summarise, the influence of psycholinguistics can inform some aspects of how we hear and process musical information. However, listening to music seems to involve multiple cognitive frameworks and individuals may shift their strategies throughout the piece according to the dictates of the listening environment. Studying musical scores and an analysis of the written notes may direct our listening; it may direct us to what we 'hear' and it may increase our enjoyment of the piece but we may do this without ever fully grasping the musical grammar or the 'encoded message' from the composer. We cannot take part in a linguistic event without sharing a common

framework of language and meaning, but music appears to have multiple voices and individuals appear to have multiple ears.

Language conveys meaning. Specific components of language convey specific meanings. Specific words are referents to external objects; they are *signs* of that object. If an individual learns the rules and signs he or she may understand another individual. The influence of psycholinguistics does however create some concerns when viewed from an aesthetic perspective. The question raised by aesthetics asks whether or not music can convey meaning in the same way that language can. Can specific components within the music be said to be referents (signs?) to the extramusical world of concepts, actions and emotions? This question is explored in the next section, however the question of how many similarities and differences there are between language and music remains unresolved.

1.2.3 Meaning in music

There are three main perspectives on the meaning of music. Firstly, the *referentialist* perspective suggests that the meaning of music lies in ideas beyond the music itself. (e.g. moods and pictorial representations). Secondly, the *formalist* perspective suggests that meaning arises only from within the work of art itself. Third, the *expressionist* positions suggest that meaning is derived from internal qualities that in turn react with external referents. There is some overlap between the positions. A brief review is detailed here.

The referentialist view suggests that the meaning and value of a work of art exists beyond the work itself. In order to discover the meaning of the music the listener must look beyond the music to external ideas, emotions and events in the real world. The function of the music, (or art work) is to inform, remind or assist you to re-experience or review and further understand something, which is beyond the music. A referentialist will argue that factors from the real world (experiences, history, education, culture etc.) act upon the artist and shape the choices they make in creating the artistic product. Each of these individual factors is a further clue to the meaning of the work beyond the music itself. Referentialists argue that music should portray an emotion and that the emotion must be specific and must serve a non-artistic end, such as nationalistic pride. Some referentialists have different views about the positive and

negative affectivity of different emotions and therefore the value of a work of art can only be judged according to subject matter beyond the work itself and the type of emotion it engenders.

The referentialist view is well represented by Cooke (1959). Cooke regards music as a language and therefore he suggests that the meanings of the terms used in the language can be specified. In his book, *The Language of Music*, he reviewed a large number of classical musical works and argued that very specific meanings exist for both individual intervals (minor second equals spiritless anguish; minor third equals tragedy) and for musical phrases (ascending minor scale equals outgoing feeling of pain, sorrow or a complaint against misfortune). Accordingly, the listener must first learn all these individual specific meanings and only then will he or she be able to understand the full meaning of the music. There is obviously some music that is deliberate in its representation of physical images (e.g. tone poems), but giving music a meaning in this way does not preclude it also having an absolutist meaning. Some composers specifically intended there to be a number of layers of meaning within one particular piece.

Cooke can be regarded as representing an extreme referentialist perspective, yet there are a number of components that exist in many music education programs, which come from a referentialist perspective. Many educators set works of art within a story or assign moods to certain pieces whilst therapists encourage the expression of repressed emotions through music. Studying music is said to be good for self-discipline whilst some forms of music are discouraged because they carry either a political or anti social message. In these and many other ways, music can be seen as moral and beneficial or immoral and questionable. In the referentialist view, all are non-musical and refer to ends beyond the music itself.

The *formalist* perspective argues that music has meaning solely within itself. The meaning and value of a work of art lies within its internal qualities; that is the same qualities that a referentialist would argue are actually signs, signifiers or representations of other external meanings.

Formalists, as represented by authors such as Hanslick (1986), assert that music has no meaning beyond itself. The meaning of art is like no other human experience and therefore the true meaning cannot be explained beyond the recognition and appreciation of the form the art takes.

“If now we ask what it is that should be expressed by means of this tone material, the answer is musical ideas... the content of music is tonally moving forms” (p.29)

Formalists accept that art works may sometimes contain references to features beyond the artwork itself, but they argue that this reference is meaningless and irrelevant to understanding the true meaning of the artwork. Formalists argue that of all the arts, music is the finest example of artistic meaning because it contains the least reference to external features. The beauty of music is specifically musical and extra-artistic meanings contribute nothing to the significance of the artwork.

Formalism is regarded by some, as being elitist. The formalist approach regards the majority of individuals as insensitive people and the true meaning of art is available on a limited basis to only a select few.

Hanslick represents the extreme pole of formalist ideas and approaches to art. However, as with Cooke and the referentialist movement, a number of claims and ideas within music education have a formalist perspective. A considerable number of music education courses largely concentrate on isolating individual formal elements of music and exploring their use or development. Likewise, many have a clear focus on the development of specific skills. Furthermore, the history of music education has continually shown a tendency to concentrate on teaching the most talented children. Even though some change has occurred in recent years, the status of outstanding and highly talented general class music teachers is still well below that afforded to those involved with serious music teaching. This is essentially a formalist perspective.

Reimer (1989), summarising formalist and referentialist perspectives concludes that:

“... each view contains some truth, each also contains major falsehoods that prevent their use as a basis for a philosophy. Somehow their contributions to understanding must be preserved while their limitations are overcome.” (p.26)

The third perspective is *Expressionism*. This perspective can be said to include elements of both formalism and referentialism and yet not simply be a combination of the two. Expressionists suggest that the meaning of a work of art is to be found within the internal formal qualities of the work. However, the external artistic and cultural influences surrounding the work of art also influence the meaning because they are a part of the internal processes of the observer/ audience member. Direct external references, such as words and content can be a powerful influence upon the artistic experience but these influences are transformed and transcended by the internal artistic form. Artistic meaning and value are frequently above and beyond the individual external referent. Thus a trivial reference to something beyond the artwork, can be transformed into something profound by both the internal qualities of the artwork and the inter-reaction between the internal and external qualities. An example of this would be the profound result of a trivial reference to a simple peasant folk dance in Beethoven's sixth symphony or Vaughan Williams fifth symphony.

Aiello (1994) takes a broader perspective on musical meaning that in some ways adopts some of all three philosophical positions. She suggests that meaning may be derived from three main areas. The first area is an intellectual appreciation of musical elements. This could include the listener's prior knowledge of biographical details of the composers life which are pertinent to the piece; appreciation of the composer / performer skill or a professional instrumentalist listening specifically to how their instrument is written for. The second area is the emotional or aesthetic response. This could include many types of response from quiet appreciation to tears and ultimately to more overt physical responses ranging from dancing through to nationalistic uprising. The third area is the association of a piece with a specific event external to the music. This last area could include music that has become associated with a particular set of words; music that has become tied to a specific event in either a personal or broader sense, (e.g. the Wedding March) and music that presents an imitative aspect of an external sound (e.g. bird song).

Each of these areas could equally be divided up into a number of subsections. Aiello further notes that the musical meanings within a specific piece are constantly available to us in all of these areas, but we are limited by our ability to focus only on

one area at a time. In this way, she asserts, we are always “correct” when we listen to music because no exclusive meaning can be attributed to one piece.

Beardsley (1981) argues that the terms “expression” in music and “emotion” in music are often regarded as being synonymous. He suggests that music is often attributed with meanings and emotions that we cannot verify. We may say that a piece of music expresses joy, but have seldom any way of knowing if the composer was feeling joyous at that moment. He concludes that if what music expresses is discoverable from listening to music, then expression is certainly not the same as the venting of emotion. Essentially, he argues that when listeners claim that a piece of music expresses “joy”, they are in fact stating that the music makes them feel the emotion of “joy”. That is, they are speaking about their own state of mind and not the meaning of or in the music. He argues that the word “expresses” should be substituted by the word “arouses”. Music *arouses* feelings but does not *express* them.

Some studies have attempted to explore whether musical meaning can be conveyed by attributing human qualities to specific musical components, (Cooke, 1959; Sloboda 1991). Sloboda (1991) carried out a study in which listeners were asked to note their own physical responses to music (e.g. shivers down the spine; tears etc.) and the exact place in the music where this occurred. It was suggested, as a result of this research, that certain musical events could arouse certain physical responses. Tears were induced by melodic appoggiaturas, sequences and harmonic movement through the cycle of fifths to the tonic. Shivers appeared to be associated with sudden changes in harmony whilst heart beat increase resulted from acceleration and syncopation. A number of participants selected the same pieces of music for noted effects (52 from 165) whilst others showed great diversity.

There are a number of difficulties with experiments that attempt to explore this. Firstly, it is difficult to show that participants are actually hearing and responding to the same musical qualities, even though they may note the music represents the same emotions. Secondly, the detailed instructions and the format of the written response may be directing the subject’s listening to specific musical qualities they would otherwise not hear and third, communication between the experimenter and the

subject is mediated by language, which may again direct or influence the subject's attention.

“ For where there is no language but metaphor in which to describe the music, there is bound to be some deviation in the understanding of metaphors, even if all subjects are are in fact hearing the same human quality” (p.329)

One further perspective is given in the *Signification* theory of musical meaning. This perspective suggests that music does have a referential relation to things beyond itself but this is best explained through semiotic terms; that is musical meaning can be analysed using the concept of sign (Langer, 1953, 1979; Gehring, 1910). Although some authors prefer to use the term ‘symbol’, essentially ‘sign’ denotes any object or event that stands for something else. Music is a sign that stands for psychological processes.

“It “articulates” or “elucidates” the mental life of man, and it does so by presenting auditory equivalents of some structural or kinetic aspects of that life” (Beardsley, 1981. p.333)

There are a number of difficulties associated with this view. First, similarity is not enough to imply sign. Mental processes may be similar to aspects of music but are not necessarily a sign of them. For example, a physical area or plan is similar to a map but does not signify it. Secondly, aspects of music can become signifiers when assigned verbal labels, or when attributed a conventional meaning by a group of individuals, for example as a code. However, music in general does not fall into either of these two categories. The third difficulty is that it is again difficult to establish any causal connection between the specific states of mind of the composer and particular fragments of the music.

Aesthetic and philosophical theories search beyond music for meanings because the listener often does not understand the significance of a musical piece. Aiello (1994) and Beardsley (1981) agree that not understanding the meaning of a piece of music occurs when the listener fails to hear all that there is to hear. To grasp the meaning of a piece of music is to organise its sounds; recognise basic sequences in different guises; grasp its use of rhythm and melody; perceive the kinetic qualities and then to perceive the human qualities within it. To accomplish this requires multiple hearings. Each additional hearing allows additional aspects of the music take shape.

One major contribution to the debate on musical meaning has been made by Meyer (1956). Meyer made two initial arguments, firstly that the various aesthetic positions can and do co-exist. He argued, for example that it is possible for an individual to believe that ‘meaning’ in music originates within the perception of the patternings within the music. It is a further possibility that these patternings may also ‘refer’ to feelings and emotions that are extramusical; that is both absolutist and expressionist. Secondly, Meyer argues that in the aesthetic debate, all sides tend to concentrate on the strengths of their own position and the weaknesses of their opponents position. As a result of this, the fundamental question common to all aesthetic positions on meaning has been neglected. Basically, they avoid any explanation as to the “manner in which an abstract, non-referential succession of tones becomes meaningful” (p. 4). In failing to grasp this basic question, proponents of all camps have in fact remained opponents of each other rather than realising the common ground they share.

Writing in 1956, Meyer is critical of previous approaches taken by psychology. Meyer argues that their search for a theory of musical meaning has consistently been based on three wrong assumptions. The first wrong assumption he terms *hedonism*: this is the confusion between aesthetic experiences and pleasurable sensations. According to Langer (1979), this creates “a conception of art as the satisfaction of taste” (p.211) and Meyer agrees with Langer that confusing a pleasurable response with an emotional response (which may or may not be pleasurable), has added nothing to our understanding of musical meaning but has simply given us “a number of like/dislike measures and a number of observations on musical tastes” (p.211). The second wrong assumption he terms *universalism*: this is the belief that responses to music are universal and good for all times, places, cultures and peoples. The third wrong assumption he terms *atomism*. This he sees as the attempt “to explain and understand music as a succession of separable discrete sounds and sound complexes” (p.5).

With these criticisms, Meyer and Langer are attempting to clear the ground by highlighting the lack of real progress made towards a theory of musical meaning by each of the previous approaches. Meyer points out that whilst it is generally accepted that an emotional response to music does take place in the listener, very little is

known about the exact relationship between the response and the stimulus. He presents a summary of the main difficulties experienced when making an attempt to study this issue. Meyer suggests that there are three areas of evidence for the relationship between a musical stimulus and an emotional response. Firstly there is the subjective evidence of individual listeners, composers and performers. Meyer cites three problems with this evidence. Firstly, it is imprecise. Musical is temporal and attempting to pinpoint the precise moment where an emotion is aroused whilst listening is very difficult. Similarly, retrospective accounts involving language have not been very accurate. Secondly, objective evidence is available in the form of behavioural responses. Although avoiding the problems of verbal descriptions, observing behavioural changes poses another set of problems. A number of these have been discussed in a previous section, however Meyer adds two further problems to this list. The fact is that one common emotion may result in a number of highly varied observable behaviours. He quotes the fact that high levels of stress can result in either total static and rigid reactions or frenzied activity, depending on the individual. Meyer also points out that observable behaviour to emotional stimulus becomes standardised through socialisation. That is, behavioural patterns that are socially expected or appropriate for the circumstances. For example, behavioural reactions within the classical concert hall are different from those in a heavy metal concert.

The third area pinpointed by Meyer is that of physiological reactions. Firstly, he points out that no accurate relationship can be found between the character or pattern of the music and a particular physiological response. Secondly, the listener brings to the situation certain beliefs about the emotional responses immediately ahead. Before the music is played, the listener has already activated a number of predispositions to certain emotional responses. Therefore it is unclear as to whether or not physiological responses are a result of the listeners' mental functions or to the actual tones themselves. Lastly, as with behaviour, physiological responses can become standardised.

All this leads Meyer to propose that understanding music can be explained in terms of Gestalt psychology. Understanding is not simply the perception of single sounds or sound combinations in isolation, but is “ rather a matter of grouping stimuli into

patterns and relating these patterns to one another” (p.6). Meyer further argues that a psychological theory of emotion is required because other available evidence from the introspections of musicians, aestheticians and listeners did not yield reliable information.

He therefore proposed his theory of “tendency and expectation”, namely:

“ Emotion or affect is aroused when a tendency to respond is neglected” (p.14)

A “ tendency” according to Meyer, is a pattern reaction that consists of a series of mental or motor responses that once activated, follow a pre ordained course unless inhibited or blocked in some way. Meyer pointed out further that the pattern reaction was equally temporal and structural; that is equally susceptible to changes or upset in the timing or the order of events. “Tendencies” are largely unconscious and occur frequently unbeknown to the individual. “Tendencies” become conscious when the reaction pattern is disturbed or the finality of the tendency is delayed or inhibited.

With specific reference to music, Meyer argues that music can be seen as a set of sound successions common to a culture, a style or a particular work. The listener establishes the nature of those patterns as ‘norms’. Norms lead to expectations, that is the listener has the expectation that the pattern will follow a set course. Meyer contends that when the course of those musical patterns is altered, inhibited or the resolution suspended, then arousal takes places and a conscious affective reaction is experienced. Meyer agrees with the position taken by Mead (1934) and Cohen (1944), that musical meaning is the result of a ‘triadic’ relationship. The three components of this relationship are the stimulus, that to which the stimulus points (or is a sign of) and lastly the conscious observer.

Embodied musical meaning is a product of expectation. As Meyer states:

“ If, on the basis of past experience, a present stimulus leads us to expect a more or less definite consequent musical event, then that stimulus has meaning. From this it follows that a stimulus or gesture which does not point to or arouse expectations of a subsequent musical event or consequent is meaningless” (p.35)

Lastly, Meyer points out that communication between two individuals only takes place when the stimulus has the same meaning for the individual who makes it and the individual who responds to it. There is therefore a necessity firstly for a set of common gestures or signs within a social group and secondly for a shared set of common responses to those gestures or signs. Without this, no communication may take place. Meyer argues that communication depends upon style, which he sees as :

“....a replication of patterning, whether in human behaviour or in the artefacts produced by human behaviour, that results from a series of choices made within some set of artefacts”
(Meyer, 1989. p.3)

To summarise, listening to, or hearing music is an everyday activity for most individuals. Often it is an activity that is entered into without much thought. Individuals respond to it and can be affected by it, and yet can often be unaware of major parts of it. Listening to music can be a simple activity that the listener engages in at an almost subconscious level. Yet attempting to discover exactly what the listener is hearing, reacting to or understanding is a complex and problematic area and results from studies that attempt to explore this area of psychological activity should be viewed and interpreted within the context of the studies reported briefly here.

1.3 The present studies

The rationale for the present study was stated earlier, however there are two further points to be noted. Sloboda and Davidson (1996) have explored two particular myths. One is the general idea that high levels of musical ability are a rare occurrence, whilst the second myth assumes that high levels of musical skills in adults are predicated on early musical attributes or capacities. Contrary to these assumptions, Sloboda and Davidson suggest that:

“.....there is evidence that much learning about music and its structures is something that takes place normally as a result of exposure to the musical products of the culture.” (p.178)

They further note that final musical performance ability may well be built upon a very common heritage rather than a rare set of individual characteristics and Sloboda (1989) suggests that many children come to performance in music through a rich and positive listening experience.

Any work that informs and therefore improves the quality of those listening experiences may considerably broaden and influence the quality of musical experiences that children come to enjoy in later life. Likewise, Sloboda (1985) has argued that many adults achieve quite acceptable levels of musical skill simply through a process of 'enculturation': absorbing the rules and the role of musical elements simply by their everyday exposure to them. Today, schools in the U.K. are suffering from an over crowded timetable with more and ever increasing demands being made upon teachers. It is therefore vital to question further the role and teaching of music in the school curriculum. If Sloboda is correct, then most children are going to acquire a full and adequate range of musical skills simply through contact with the musical products of their culture; something which they are surrounded by in ever increasing amounts. We therefore need to constantly examine the role of music education and the role of listening to music within that education in order to augment and not simply replicate that which children are possibly able to acquire from their everyday contact with culture.

It is perhaps interesting to note that the difference between what pupils learn formally and informally within / outside school occurs in many subjects, yet it tends not to be questioned in other subjects. Whilst there is no definitive answer to this issue, there are a number of possible explanations. Firstly, schools are already legally bound by what many see as an overcrowded curriculum. The true value of subjects beyond the essential core curriculum, such as music and art, is often questioned far more than other subjects. Eisner (2001) suggests that, "The general public does not think of music as the product of complex forms of thinking. In terms of educational priorities music is regarded as nice, but not necessary....Music begins to become important when the public believes it contributes to extramusical outcomes, such as its highly touted contribution to spatial reasoning and math performance, for example." (p.5). Secondly, music is one activity outside school, in which a relatively large number of pupils are actively and prominently involved. For example, many pupils take private instrumental lessons and this is accepted as normal, whereas relatively very few pupils take private art lessons. Many educators see learning instruments as 'the true music education'. They also see that this requires highly skilled teachers and therefore many teachers view music as a subject which should be taught outside school, for

those who wish it, with the teachers who have the skills to do it. Therefore a large number of teachers question their own ability and their own individual contribution to the music education of the children in their care. The whole issue of teachers / parents view of music education and influences which in turn lead music teachers to feel it necessary to defend their subject and its place in the curriculum so strongly is very interesting and would certainly be worthy of further research.

Secondly, Rodriguez and Webster (1997) state:

“...a developmental view of affect is pedagogically useful: the determination of response tendencies to musical sound at any given age can assist in improving music teaching and learning at that age. Second, a developmental view of emerging cognitive skills is also desirable: simply asking children what and how they think about music reveals strategies for musical understanding and valuing” (p.9)

The present study makes an attempt to further explore a number of those “determinants of response tendencies” and contribute to a “developmental view of emerging cognitive skills” by further exploring the development of style sensitivity in children aged between 4 years and 16 years old as originally operationalised by Gardner. In his original study, Gardner (1973) was attempting to discover any age related stages / changes which may be applicable to a view of the development of style sensitivity in children. Castell (1983) suggested that learning and some degree of enculturation also played a considerable part in children’s’ perceived ability to be sensitive to musical style. Lastly, Hargreaves and North (1997) spoke of the myriad of social influences impacting on the cognitive development of the individual. In looking at the development of style sensitivity, it is necessary to be aware of the individual role played by cognitive development (see also discussion of this by White (1997) in section 3.1); that which is learned through formal and informal education (education and enculturation) and the social influences which affect how the combined effect of these individual elements is demonstrated.

The studies in this thesis build upon and extend the previous reported studies into the development of style sensitivity and attempt to unpick a number of social, educational and developmental influences that may affect the perceived ability of children to be sensitive to musical styles. Chapter 2 sets the study within the context of

developmental and social psychology. In the first section, two theories of development are contrasted; the stage theory of development as outlined by Piaget and the theory of artistic development as outlined by Gardner. In addition to this, a number of developmental theories of musical development are outlined. The third section explores the social psychology perspective. Firstly, it is argued that music is a social product and can be subjected to sociological enquiry. Secondly, there is a brief survey of recent developments in both the social psychology of music and also developmental social psychology.

Chapter 3 presents the more immediate background literature to this thesis firstly by presenting a more detailed picture of Gardner's theory of artistic development and secondly by presenting the details of the four previous studies into the development of style sensitivity. A review is given of each of the studies done by Gardner (1973a), Castell (1982, 1983), Tafuri et al. (1994) and Hargreaves and North (1999). Details are given of methodological differences between the studies and a summary of the contrasting results obtained in each study.

Chapter 4 reports on a number of methodological issues. A number of previous studies that have explored possible sources of influence or bias in test situations, such as experimenter expectations and test anxiety are reported. During the initial design of the experimental procedure employed in experiments one to four of this present study, the findings of each of these studies were noted and some changes were made to the overall test design in an attempt to reduce any experimenter distortion. The last section of Chapter 4 presents further details of the experimental procedure used in experiments one to four of this present thesis. Details of the selection and recording of musical examples, the form of the response sheet and the procedure used in marking and analysing participants' responses are also included.

Chapter 5 presents the first experiment which examines the importance of one social context, namely the context in which the test is operated. In this experiment half the participants receive the style sensitivity test within the context of a music lesson and the results are contrasted with responses from participants in a control group. The

experiment explores whether or not participants display a greater tolerance and willingness to listen to a wider variety of musical styles within that environment than participants in a non-musical context. Chapter 6 presents the second experiment in which the same style sensitivity test is operated using three different test presenters. A number of experimenter attributes are varied. The presenters varied in age with two being much younger and closer in age to the test participants, and also two of the presenters were male and one was female. The experiment sought to explore whether or not the two younger presenters could add prestige and increase participants' tolerance of all or some of the test material. Chapter 7 details the third experiment which contrasted two sets of responses from participants in different areas of the United States with a third set of responses from participants in the United Kingdom. The purpose of this study was to contrast the responses between participants who experienced different music education programmes.

Chapter 8 presents the results of a further experiment in which the introduction to the test was manipulated. In the first condition, an attempt was made to increase the participants' motivation to do well in the test, whilst the second condition attempted to make the test more competitive. The purpose of the experiment was to explore whether or not participants increased motivation would increase their tolerance of the test material and therefore perform more accurately. Chapter 9 presents the last experiment in which the musical material used in the test is manipulated. In the first condition, a broad grained but effective method of measuring the level of contrast between the classical and popular musical samples is developed and employed in a new set of musical examples. In the second condition, the chronological time period between the four classical eras is increased in order to further increase the level of contrast between the classical styles. The purpose of the experiment is to explore the effects on subject responses of manipulating the music material in a systematic way. The experiment further explores the contrasting results obtained Gardner (1972a) and Tafuri et al. (1994) with respect to participants' responses to classical samples from adjacent and non-adjacent eras.

Chapter 10 reports on an exploratory study carried out to develop a new form of test procedure. The aim of the procedure was to enable further studies to explore the

development of style sensitivity in children below the age of six years old. Gardner (1972a) had commented on the high level of accuracy achieved by “even his youngest participants” who were aged six years. No published study of the development of style sensitivity in children has ever incorporated participants under the age of six years old. The reason for this was probably that the test procedure incorporated by all the previous studies in this area is unsuitable for most young participants. The procedure is long and required participants to record their written response on a response sheet and therefore the effectiveness of a new test procedure was investigated in order that the development of style sensitivity could be further explored in much younger children. The chapter presents a brief review of some recent literature on the musical competencies of very young children and also details some previous research incorporating a number of testing procedures more suited to younger children. This previous literature influences the design of the test procedure reported in this thesis, and an account of the results from this exploratory study using the new procedure is subsequently given.

Chapter 11 presents a review of the experiments carried out in this thesis and in the second section, a model for the development of style sensitivity is proposed. In this model, a distinction is made between the terms ‘style sensitivity’ and ‘stylistic competency’ and the results of the experiments carried out and reported in this thesis are set within the context of this model. The final sections detail a number of implications for future research and also for education.

Chapter Two

The developmental and social psychology of music

"Tuesday night I reorganize my record collection. I often do this at periods of emotional stress... I try to remember the order I bought them in: that way I hope to write my own autobiography, without having to do anything like pick up a pen... .. and when I've finished, I'm flushed with a sense of self, because this, after all, is who I am"

(Hornby, 1995. p.54)

The experiments in this thesis draw on two main areas of psychology, namely developmental psychology and social psychology. Developmental psychology explores the appearance of qualitatively different age related stages whilst social psychology explores the interactions between social groups and between individuals; the formation of social attitudes and the interaction between social influences and genetic / cognitive processes within the individual.

This chapter contains three sections. The first section reviews three perspectives of developmental psychology; the second section gives a brief review of the social psychology of music whilst the third section outlines a number of developmental theories which apply specifically to music. In section one, in order to set the work within a historical perspective, a brief review is given of the work of Freud. Freud regarded children simply as mini adults and therefore his theories of how development from childhood to adulthood occurred were based solely upon his work with adults. This brief description of the work of Freud is followed by an outline of Piaget's cognitive stage theory of development. Piaget's stage model of child development concentrates mainly on the cognitive aspect of development and therefore is an interesting contrast to the theories of Freud which were placed mainly within the affective domain. This is followed by a brief introduction into the theory of artistic development by Gardner. Gardner (1973b, 1994) draws on the pioneering work of both Freud and Piaget and then proceeds to argue for a comprehensive role for the arts within development psychology. Only a brief outline of Gardner's work is given in this chapter. A fuller description of his theory of artistic development

provides the immediate background into the research into style sensitivity, and therefore this is presented in a subsequent chapter.

In the second section, there is a brief critique of the main theories that have attempted to apply developmental psychology specifically to the development of musical skills. This is followed by an outline of the ideas and terminology used by two writers, namely Koopman (1995) and Koroscik (1997). Koopman proposed a model of musical development which could be said to encompass three different and disparate stage theories of musical development proposed by Swanwick and Tillman (1986), Gardner (1973b) and Gardner, Phelps and Wolf, (1990). Koroscik, on the other hand, attempted to describe how children might come to understand works of art through description of, and reflection upon her own personal experiences. Both the ideas and the terminology of both these writers are adapted and modified in the final chapter of this thesis when a possible model of style sensitivity is proposed.

In the last section a number of sociological issues are presented and discussed. It is argued that music can now be viewed as a social product and therefore a valid subject for sociological inquiry. The increase and the changes in the role of music in everyday life have changed rapidly within the last few years. Therefore investigating such a social product outside the social environment in which it is heard, seems to be no longer a desirable or accurate activity

2. 1 Theories of developmental psychology

Freud began his interest in the human mind by studying cases of hysteria. His primary assumption was that adult hysteria was the product of a psychical trauma that had been forgotten by the patient. His treatment consisted of hypnosis which enabled the subject to remember the trauma accompanied by more suitable, pleasant and appropriate emotions. Freud developed a number of further ideas including free association and his total theory of psychoanalysis. Freud was the first person to suggest that the mind consisted of a number of dynamic mental forces some of which were conscious and some unconscious. At times they operated in harmony whilst at other times they operated in opposition to each other (Freud, 1972).

However, in terms of developmentalism, Freud's theories had a number of limitations. Firstly, Freud had paid scant attention to the actual study of children. He spent a number of years studying children with cerebral palsies, but this was a brief interest and did not involve the study of any developmental processes. Freud tended to regard children as small adults. Certainly in matters of cognition, he suggested that children and uneducated adults often behaved in similar ways. He regarded children as having the ability to conceptualise very complex notions. Young children, he argued, did not lack the reasoning ability of the adult, only the large amount of factual information. Freud (1938) also felt that:

“ The direct observations of children have the disadvantage of working upon data which are easily misunderstandable” (p.599)

Secondly, Freud admitted that his perspective of psychoanalysis forced him to concentrate on the affective and subconscious aspects of the personality rather than the logical, conscious and cognitive elements. Therefore areas such as language development, logical reasoning and problem solving were either ignored or seen as defence or substitute behaviours. Thirdly, Freud made a large number of claims regarding children that were little more than speculation. His view that new born infants hallucinate about the mother's breast or that young children are preoccupied with ideas of marriage and sex, were totally unexplored and unsupported. In order to develop his ideas, Freud did not begin with the infant and explore how development occurred in any area; rather he often began with the end product of a neurotic adult. Firstly, he would describe how this adult differed from a healthy adult personality and follow this with a retrospective analysis of the individual life to establish which experiences had become fixation points in the adult. For Freud, no experience or emotion changed over time and memories existed unchanged throughout life. Events occurring later in life did not react with and change the nature of an earlier experience.

As Gardner (1973b) stated:

" The development of the human personality represented for Freud, a series of layers deposited one upon another, each stratum retaining its specific character" (p.4)

As such, the Freudian perspective saw the child as simply a more immature 'mini adult' and therefore there was no requirement to investigate human development in terms of the reorganisation of mental structures. The end state of human development was defined as the "healthy adult human personality" and this was achieved by the child passing through a series of psychosexual stages, all with an affective element. At any point certain experiences may create fixations and points of regression with some affective elements remaining throughout adult life.

In contrast to Freud's mainly affective and adult centred theory, Piaget proposed an alternative theory of child development based on a number of qualitatively different cognitive stages. Many writers (Wadsworth, 1996; Elkind, 1976 and Gallagher and Reid, 1981) have fully reviewed the work of Piaget therefore only a brief synopsis is given here. According to Piaget's theory, the infant begins to experience the world through the sensory and the motor systems. By the age of 2 years the normal healthy child will have begun to talk (symbolic representation), and show an awareness of the affective (feelings interacting with intentional actions). Between the ages of 2 and 7 years Piaget suggested the stage of pre-operational thought. During this stage the main advances are in the use of symbol systems: there is a reduction in the child acting primarily in a sensorimotor mode and an increase in the ability to represent actions internally. This stage is further characterised by the development of symbolic play, egocentrism and inability to conserve volume. The third stage of concrete operations takes place between the ages of 7 and 11 years old. This stage is characterised by a developing ability to conserve volume, less egocentricity and an ability to reverse mentally an observed chain of actions. Piaget's last stage is known as formal operations, which begins around the ages of 11 or 12 years old. During this stage, cognitive potential for reasoning reaches its maximum potential. This stage is characterised by the ability to reason in a hypothetical and deductive way, the ability to construct new knowledge based on internal reflection of available knowledge and the emergence of idealistic feelings.

Piaget's stage theory has received some general criticisms and these include the use of open ended and subjective informal interviews which lack validity and reliability; tasks have an over dependence on verbal skills (Donaldson, 1978); lack of control

task alongside each experiment (Bryant, 1974); and the use of terminology that is too vague to be of any real use (Brown, 1983). Gardner (1973b) expressed a number of further criticisms of Piaget's stage theory. Firstly Gardner argued that Piaget did not fully address the role and the development of the affective domain and secondly, he argued that Piaget's 'end state' of the developmental process was that of the logical and scientific thinker and this term was an inadequate description of the fully developed human personality; it described only one facet of this personality. Gardner further argued that full participation in the arts was possible long before the final Piaget stage had been passed through. Lastly, Gardner (1973b) disagreed with any developmental theory that incorporated so little of the arts.

“... “Full participation in the artistic process” as an end state seems at least as creditable as the “normal adult” personality generally cited by the Freudians and the “scientific thinker” embraced by the Piagetians” (p.23)

Piaget argued that the affective and cognitive elements of human development were both combined and linked but also separate. In one way, Piaget saw the affective as the underlying energy that drove the more logical realm of thought. The cognitive and affective were clearly linked and dependant on each other; the affective providing parameters such as the initiation and the strength and length of reaction; whilst the cognitive element shapes and structures the response in a logical way. As Gardner states; “ Piaget depicts affect as a motor for all behaviour, it determines the force of action but has no structure” (p.6). On other occasions, Piaget suggests that the affective realm has its own line of development, mainly associated with social objects: “having its own unique history, emphases, schemes, reactions and energy” (p.6).

This imbalance in the emphasis between the cognitive and the affective elements of the personality is often seen as a major criticism of Piaget's work and the lack of recognition of the social and cultural factors in intellectual development has, in recent years, directed many towards the work of Vygotsky. Wadsworth (1996) however, has made a brave effort to redress this balance. Wadsworth argues that in fact Piaget recognised the importance of affectivity from his earliest writings and that he never intended the imbalance in his own work to be seen as a value judgement: Piaget did

not view the affective domain as being either secondary or inferior to the cognitive domain. Secondly he suggests that Piaget made a conscious effort to establish a framework for what he saw as the most manageable questions first; that is, those questions relating to the cognitive development. Thirdly, Wadsworth argues that affectivity is often regarded as highly subjective and “intellectual mushiness” by the scientific community. It is therefore the scientific community of psychologists and educators who have tended to afford cognitive development a higher status than affective development. Piaget placed affectivity at the centre of intellectual development. Wadsworth argues that it is academics who have subsequently brought about the greater imbalance by directing almost all of their research towards the cognitive aspects of Piaget’s writing. Piaget always insisted that affectivity played a key role in the individual’s decisions as to which concrete experiences were selected for the construction of schema. Wadsworth has argued that this Piagetian perspective has, in recent years been constantly ignored or forgotten (Brown and Weiss, 1987).

To summarise, Piaget’s theories relied as heavily upon the cognitive development of children as much as Freud’s theories relied on the affective side. Piaget argued for a distinct set of stages whereas Freud described a number of sequences that blended into one another. Piaget saw cognitive stages that were qualitatively different whereas Freud described a sequence of dominant affects that may reappear totally unchanged throughout the life of the individual.

It was these two dissimilar models that led Gardner to postulate his theory that an integrated view of human development was achieved through those activities, which combined cognitive and affective elements. Gardner argued that those activities existed within the arts. He therefore proposed that:

" the artist, audience member, performer and critic provide a viable and holistic end state for human development" (p.7)

Gardner argued strongly therefore for the arts to be included in developmental psychology. He argued that full participation in the arts was an acceptable “end state” of development. As a result, Gardner developed his own developmental model involving three systems: namely, the making system, the perceiving system and the

feeling system. Criticism of Gardner's approach and a more detailed account of his three-system model is given in a subsequent chapter.

2.2 Developmental theories of music

In spite of the criticisms, Piaget's developmental 'stage' model has had a major impact on subsequent developmental theories of the arts and music. The idea of development through qualitatively distinct stages has influenced the model of musical development put forward by Swanwick and Tillman (1986) and also the developmental model of aesthetic development put forward by Parsons (1987). Furthermore, a large body of research has been done on the idea of musical conservation. This research adopts the Piagetian concept of conservation of matter and applies it to musical material and was pioneered by Pfleiderer (1964).

The first theoretical model by Swanwick and Tillman (1986) was based on their analysis of 745 musical compositions produced by children of different ages. These authors proposed a spiral model of musical development, suggesting there were distinct age related trends. Swanwick and Tillman suggest that there are four levels of operation namely:

Mastery (0-4 yrs.) - child learns to control sounds.

Imitation – (4-9 yrs.) child makes representations or illustration of the world through imitation.

Imaginative play (10-15) – child makes a creative contribution rather than just imitation.

Metacognition (15+) – child becomes more aware of its own musical ideas.

Each of these levels, they suggest display specific musical phenomena. They group these phenomena under the titles ***Materials, Expression, Form*** and ***Value*** respectively. These headings relate to the musical phenomena that are most prominent at each level. The four levels are lastly applied to another level of discrimination in that each level develops from left to right; the left hand side being related to more personal aspects whilst the right hand side moves towards forms of social sharing.

Hargreaves and Zimmerman (1992) have expressed two concerns about Swanwick and Tillman's model. The first concern surrounds spiral and stage models in general. Swanwick and Tillman's levels essentially represent a developmental stage in the same sense as Piaget's stages; additionally Swanwick and Tillman suggest a move from personal towards social within each stage. Stage development theories are essentially all context free explanations of developmental change. They do not take account of the interaction between environmental experiences and objects and the role these may play in the moulding thought. Secondly, the raw data used in the study, that is the compositions collected can only be said to "describe or illustrate the model, rather than provide a deductive test of it" (p.381).

One final point is that essentially Swanwick and Tillman propose that their model is one of "Musical Development" and yet the data used to propose the spiral model consisted only of musical compositions. This rests on the assumption that the developmental skills, processes and stages involved in children's compositions are identical to those in other areas of musical skills such as listening, performing and representation. Other research (Davidson, Howe and Sloboda, 1997; Geake, 1999) suggests this may not be the case. The issue of Swanwick and Tillman's model and its relationship to Gardner's ideas, is discussed in a later section.

Pflederer Zimmerman explored the application of Piaget's concept of conservation in children's musical development. In this work, she hypothesised that the ability to appreciate certain aspects of music, e.g. musical transposition or identification of a melody played at two different speeds, depended on the ability to 'conserve' elements of the music and therefore would not appear until the child had reached Piaget's stage of concrete operations. The results of her work suggested that whilst the 8-year-old child was capable of 'conserving' certain musical properties, younger children could not do so. The presentation of stimuli to the child were modelled on Piaget's mode of questioning namely:

" You will hear this short tune played several times. Each time it is played you will hear it followed by a second tune. I would like for you to answer these questions about the two tunes: Is the second tune the same or is it
different from the first tune?
What was different about it?
Are they the same in any way?

How do you know, tell me more about what you heard"
(Pfleiderer Zimmerman and Sechrest, 1970, p.27)

The 'correct' answer from the child was defined as the ability to see that the tunes were both the same and different. If a child simply answered 'same' or 'different', then the response was considered to be demonstrating non-conservation. This response was based on the Piagetian concept of 'centration' where pre operational children 'centre' on one feature of an object (e.g. length or width where volume is defined by both); concrete operational children are able to 'decentre' and observe both.

Pfleiderer Zimmerman discovered that performance on 'Piagetian' style musical tasks did improve with age, but some authors (Serafine, 1980) were more cautious suggesting that:

"...whilst the discovery of the relationship between task performance and age is a crucial one, it is not sufficient evidence for either the validity of the tasks or the hypothesising of stages. Indeed, in almost any cognitive or perceptual task, older children can be expected to do better than younger ones" (p.9)

Pfleiderer Zimmerman proposed five conservation concepts; metrical grouping, transposition, inversion, augmentation and diminution, and identity, and argued that the acquisition of these concepts was vital to musical development.

There were a number of methodological shortcomings expressed regarding this research. (For a full summary see Hargreaves, 1986). Firstly, in the original Piaget experiment, the children saw the physical action of rolling out the clay from one shape to another. In Pfleiderer Zimmerman's experiment, no such physical comparison was available. This difference has recently been explored further using a MIDI based sequencer (Hargreaves, 1999). Results suggested that children did obtain more correct responses when they were able to observe a 'live' transformation, rather than simply hear the transformed pieces on a tape recorder. Secondly, Gardner (1973b) pointed out that analogies with Piaget were forced because any change in a musical stimulus actually does make it different. There were additional minor criticisms of the studies including original sample size and use of musical terminology.

Bamberger (1980) asked children to clap an improvised rhythmic figure and then asked participants to, "Draw a picture of your claps so you can remember them next week" (p.1). Further studies by Bamberger included similar tasks but employed children of varying age groups and also adults. She concluded that within her own work, it was possible to identify qualitatively different stages in musical development, at least with reference to rhythmic awareness. Bamberger suggested that the representations participants made of their clapping, illuminated the way in which they perceived and grouped rhythm. She found that participants tended to represent their clapped rhythm in one of two ways. The first strategy for grouping rhythm, she termed 'figural'. This strategy was related to body movement in that rhythms were collected together in 'chunks', each of these chunks relates to one body movement. The second she termed 'metric' strategy. In this strategy the focus was on the measurement of durations relative to a fixed reference point.

“..events which are the same in duration remain the same in spite of contextual function and regardless of the position of the event in a particular figure”
(introduction. p ii)

Her findings suggested that whilst the figural and metric strategies of representing rhythms could be associated with pre operational and concrete operational behaviour respectively, they were also not limited solely to those age groups. She observed that whilst it could be stated that the figural strategy was observable in and typical of the pre operational thinking of children below the age of seven years, the strategy also could be identified in adult participants. However, adults were able to adjust their thinking and understand the metric strategy following training whereas the younger children could not do so and remained firmly within the figural strategy. In addition to age, there was some evidence that musical background played some role in determining which strategy was used. Bamberger noted that the metric strategy was more characteristic of those who played musical instruments and regularly read from a musical score; whereas those participants who 'played musical instruments by ear' adopted the figural strategy.

Bamberger suggested that it was possible to identify qualitatively different stages in musical development but she did not claim her results indicated any particular support

for the Piaget theory of development. Her primary concern was that her work had identified the two separate ways of knowing and her resolve was that both these models be accepted and recognised in music education in order that they may both be taught and children encouraged to move back and forth between the two strategies.

As a further means of assessing the extent to which developmental changes in the arts mirrored the intellectual development of the Piaget theory, Machotka (1966) examined the criteria used by participants to evaluate paintings. In this research he presented individual children between the ages of 6 years and 12 years with three reproductions of paintings. Participants were asked to state which painting they liked most and which painting they liked least. Each child was also asked to give reasons for their choices. Machotka hypothesised that justifications for choice made on the grounds of subject matter and colour would precede those preference decisions made according to realism, contrast and harmony of colours. His justification was that selection by subject matter and colour would require no more than pre operational thought whereas for a child to make judgements according to realism some form of conservation would need to be present. That is, the child would need to conserve the image of the object in the painting and compare this with the image of the object gained from experience with the world and held in memory.

The results of the research suggested that style, composition and luminosity as criteria begin to appear at or around the start of adolescence. Machotka suggests that the criteria of style imply the "hypothetical existence of several manners of representation". (p.883) He concluded that:

" The observer cannot judge style...if he knows only one; he can judge it only in comparison with others which, at the time of judgement, are imagined or hypothetical" (p.884)

Gardner described the research on musical capacities in children as " scattered". He pointed out that what children knew about the arts was a very limited area of psychology and basically the research could be set within one of two frameworks. Firstly, those studies that were set in the framework of the work of Piaget, as reviewed above, and secondly, studies which tend to view the end product of

development by using older participants (Ecker, 1963). He saw the ways in which children perceived and understand the artistic process as a valuable but hitherto, neglected area of research.

Gardner, Winner and Kircher (1975) suggested that there was validity in exploring children's thoughts about the artistic process. Namely, as the child grows older, his / her role in the artistic process will most likely be as an audience member, perceiver or critic. There will be benefits from any research that informs educational practice and therefore increases the level of participation that children will be able to play in each of these particular roles. Secondly, they suggested that given the vast amounts of money spent by school systems on museum, theatre and concert visits, as well as other cultural events it seemed odd that no research existed which shed light on the extent to which children were affected by these experiences.

Lastly, Gardner suggested that the arts were an ideal medium for further research into human development:

" Indeed, participation in the arts is so natural and integral a part of human growth that an understanding of this process should provide important clues to many pivotal questions of human development" (p.23)

There was therefore a body of knowledge that supported Gardner's claim that research and exploration into the development of the arts could provide vital evidence regarding human development. The work of Pflederer Zimmerman (1967), Bamberger (1979, 1980) and Machotka (1966) although having suggested some elements for support for the Piagetian framework had raised many issues which required further investigation. It was against this background that Gardner set about his research into the ability of children to be sensitive to artistic styles.

2.3 Two developmental theories in the arts: Koopman (1995) and Koroscik (1997)

In the final chapter of this thesis, a developmental model of style sensitivity is proposed. This model draws on the terminology and ideas put forward by Koopman (1995) and Koroscik (1997). Koopman is offering a philosophical view of theories of musical development and as a consequence he utilises a number of terms which in a

strict psychological sense, may have other more distinct meanings. Therefore a brief resume of his use of terminology is presented here. Koroscik's (1997) paper attempts to explore what and how young children can understand works of art. Although her work deals exclusively with looking at paintings and is based on personal reflection, she suggests three strands of cognitive development which are adopted into the proposed model of the development of style sensitivity. Therefore a brief review of her ideas are also presented here.

Koopman (1995) has carried out a philosophical review of three stage theories of musical development; namely Gardner (1973b), Swanwick and Tillman (1986) and Gardner, Phelps and Wolf (1990). In this review he suggested a theoretical framework that established a means of combining the distinct three theories. He further argued that developmental psychologists often explore different developmental dimensions simultaneously and that they ignore "the fact that various aspects of the musical domain have different courses of development" (p.51).

In his article, Koopman uses a specific terminology to describe his theoretical framework. A brief description of his terminology is given here.

-Domain: The main and overriding area to be defined is the '*domain*'. The *domain* is the broad area of reality. *Domains* may vary in a number of ways from each other and one domain may be a part of a broader *domain*. The domain is the area in which development occurs. For example, the development under investigation in this thesis takes place within the *domain* of music; which is part of the larger aesthetic domain. As knowledge and experience are accumulated, so the domain develops and the shape of the domain is created. The shape of the domain will obviously differ between individuals depending on their musical knowledge and musical experience. In each domain, a number of *dimensions* can be distinguished.

-Dimension: The dimension is a distinct area within the domain e.g. composing, improvising or listening. Development may take place within a dimension both at a different rate and in a different way to that development occurring simultaneously in the domain.

Therefore Koopman argues that development may take place simultaneously, but at different rates, within two distinct and individual spheres.

-Horizontal reconstruction: In the *horizontal reconstruction* the development takes place within the domain. Musical experience and musical knowledge create greater musical understanding and an infinite number of further dimensions can be traced out.

-Vertical reconstruction: In the vertical reconstruction, further development within the dimension takes place. This development may be independent of , or in tandem with development taking place in the horizontal domain.

According to this analysis, development may follow different courses and may occur at different rates. Koopman is therefore suggesting that development in the aesthetic domain takes place in the horizontal plane. As a function of this development, the domain of music is further created. The shape of this domain changes according to musical experience, musical knowledge and training. The parameters of the domain may also change according to the rules of inclusion; that is individual aspects of the domain may be included or excluded as development progresses. Music can be the main domain; an aspect of the aesthetic domain or an area of the symbolic domain that also included mathematics. Within the *domain* of music, further dimensions are defined; the *dimension* of listening would be an example. Therefore, the dimension of listening may follow a different vertical developmental course occurring at a different rate to the horizontal development of the musical domain. Koopman points out that development theory allows for two forms of relationship to exist between stages. Firstly, those situations where the new stage encompasses its predecessor. He terms this 'inclusion', whilst secondly, there are those situations where the new stage substitutes for its' predecessor. That is, no trace of the earlier stage is apparent.

In the second theory, Koroscik (1997) attempts to examine her own developmental process as child with reference to the world of art. Her basic claim is that understanding works of art is multi faceted. She attempts to define three interrelated facets of cognition 'that come into play in the acquisition and transfer of art understandings' (p.146). The first strand contains the individual's knowledge base.

This is the amount of knowledge, skill and experience directly related to the artwork that the viewer has accumulated. The second strand contains the viewer's knowledge seeking strategies. Strategic knowledge is the set of cognitive steps that aids the construction of new knowledge, directs the search for new knowledge and enables the individual to apply previous knowledge to new situations. The third strand is the disposition to learning. That is the individual's willingness to learn and seek new understandings. These three facets of cognition come into play when new knowledge is acquired in one context and is transferred to another. Bransford, Sherwood, Vye and Rieser (1986) regard this transfer as the hallmark of intellectual development. Koroscik summarises:

“ For transfer to occur, a person must be able to activate the contents of his or her existing knowledge base when search strategies call for that knowledge, skill or experience. Moreover, a learner must be willing to learn” (p.146)

Koroscik argues that unless learning employs these three strategies, then effective learning does not take place. A broad knowledge base acquired through memorisation but with poor transfer skills will often forget or be unable to apply that knowledge. Similarly, providing students with a set of knowledge seeking skills without an accompanying knowledge base will also be unfruitful. So Koroscik suggests that three individual cognitive processes are involved in learning and understanding works of art. In turn, these three processes combine in one activity, namely transfer.

The terminology and theoretical models used by both Koopman and Koroscik will be amended and incorporated into the last chapter when a developmental model of musical style sensitivity is proposed.

2.4 The social psychological perspective

In this next section we will explore some of the perceived limitations and criticisms that traditional developmental theories have attracted. One result has been a call for a greater acknowledgment of the social dimension in human behaviours and development and an outline of the claim social psychologists make for this recognition is presented. We will furthermore look at a second imbalance that some social psychologists claim exists within social psychology itself. That is the discrepancy they see existing between the number of studies involving participants in

the early years and those in the later years of development with the majority of studies concentrating on adult society and exploring social relationships and negotiated meanings that are already mature and firmly established rather than in the process of developing. In the third section we review a call for a discipline of developmental social psychology and in the last section, social influence, social context and social psychology are discussed.

2.5 Historical background of social psychology

Social psychology, as it is known today, dates from the early 1900's. Although there is some debate between the American and European traditions as to exactly when it first began, ever since these earliest dates two main strands of emphasis have emerged. There are a number of ways of denoting these two strands; Hargreaves and North (1997), following Hewstone and Manstead (1995) make reference to *psychological social psychology* and *sociological social psychology*. Hewstone, Stroebe and Stephenson (1996) refer to social *psychology* and *social psychology*, (their italicising). The first emphasis is on social *psychology*, which stresses the individual and intra-individual processes, as does the parent discipline of psychology (e.g. McDougall, 1908; Simmel, 1908). The second emphasis on *social psychology*, focuses on the role of social context for individual processes (e.g. Lindner, 1871; Durkheim 1974, Ross, 1908). Social *psychology*, as represented by such writers as McDougall, focussed on the innate propensities and capacities of the human mind and how these might influence and shape the social world of humans. Durkheim, on the other hand as a *social psychologist*, criticised individuals who attempted to explain society in this way without acknowledging the social context in which development takes place.

Durkheim argued that society predated the individual. Historically, human consciousness was firstly collective with people being aware of themselves primarily as a member of a tribe or clan. According to Durkheim's view, human evolution has involved the gradual transition from this collective view over to the present day view of differentiated forms in which people think of themselves as individual, distinctive and unique. For Durkheim then, the concept of the 'individual' is a relatively recent

phenomenon. This view was obviously in conflict with the psychological view that human actions were the result of a number of fundamental structures in the mind. Durkheim held two main criticisms of this psychological viewpoint. Firstly, that no one has satisfactorily demonstrated these fundamental structures. Secondly, there are, throughout both the present and the historical world, a large variety of modes of thought. Individuals with almost identical brains to ourselves have always and still do, conceive of the world in very different ways. Even notions such as time, space and music appear to be culturally variable. Durkheim held that the individual human personality was not the irreducible element of social life, but was formed within an already existing cultural environment. Through a process of socialisation, individuals learn the beliefs, customs and conventions of their culture.

2.6 Developmental social psychology

As stated earlier, developmental psychology historically, appeared to be synonymous with child development with little attention paid to any further developments occurring throughout adult life and also to developments occurring through or influenced by social interactions. Developmentalism neglected social circumstances. Similarly, Durkin (1996) has criticised social psychology for not attending to the developmental component within social psychology,

“ (social psychologists)...rarely describe the developmental aspects of their concerns, and rarely entertain a developmental dimension to their explanations” (p.47)

Durkin emphasises that responding to others, evaluating and negotiating understandings are not phenomena that are unique to the adult world and their developmental histories are surely relevant to their mature functioning. “ We develop as social beings over a long time, and development ceases only with death” (p.47). Developmental social psychology, he suggests, should aim to explore the developmental courses and social contexts of human social behaviour. At this juncture, two further perspectives emerge within developmental social psychology. Firstly, social interaction and the growth of understanding which explores the social context in which cognition develops and secondly, understanding and the social world which explores the social content.

The first perspective explores the socio-cognitive issue. In the past, developmental psychologists frequently explored cognitive development in individual children. This was often done through the setting of a task that the child completed alone. Doise, Mugny and Perret-Clermont (1975) and Mugny, Levy and Doise (1978) repeated the Piagetian task of liquid conservation with groups of children. The results suggested that the generating of contradictory perspectives through social interaction promoted an awareness that a variety of dimensions should be taken into account in order to solve the problem. That is, the problem was solved, and learning took place through the process of verbal and social interaction between the individuals within the group. Research on older children by Light, Littleton, Messer and Joiner (1994) also suggested that 11 year olds also fared better on a complex computer problem when working in pairs. Furthermore, the advantage gained by the participants who worked in pairs was carried through into another similar task in which they were tested as individuals. Authors disagree as to the exact way in which social interaction appears to promote or influence cognitive development. Doise and Mugny (1984) suggest socio-cognitive conflict whilst Winnykamen (1990) suggests the role of imitation. Further research may suggest much simpler explanations such as the increase in confidence gained from working alongside others. However, it is likely that the ways in which social interactions influence cognitive development will turn out to be many and varied.

The second perspective explores the increase in understanding society and social knowledge. Livesley and Bromley (1973) found that children's descriptions of the characteristics of others followed broadly age related stages. Furth (1980) suggested that understanding of societal structures advanced in stages with each stage taking several years. His research centred on children's understanding of the working of a local shop. The descriptions were broadly found to follow Piagetian stages. Similar patterns were found by Emler, Ohana and Moscovici (1987) who explored children's understanding of the role of the schoolteacher. This study was carried out in two different countries; Scotland and France and whilst overall developmental trends appeared to show some similarities to Piagetian patterns, a number of cultural differences did appear between the participants from each country. Scottish participants saw their teacher as enforcing school rules out of obligation whether they

were fair or not. French children felt their teacher should do what was fair regardless of the rule. Participants appeared to share the outlooks of their respective communities suggesting that social cognition may vary with social context.

There appears therefore to be a growing body of evidence that children's views of society's structures and social relationships change and develop in "developmental stages that are broadly similar to cognitive developmental stages" (Furth, 1980). The developing individual needs to discover the rules, properties and knowledge of the social world they live in. This discovery is achieved through participation and participation may affect the content and cognitive process of learning. The challenge for developmental social psychology is to try and explain the individual roles played by the developmental changes in capacities and the social contexts in which they develop.

2.7 The social psychology of music.

In this section we now turn to the specific issue of social influence on music and how the arguments from a broad range of other disciplines have influenced and shaped the social psychology of music.

Davies (1978) makes the claim that psychological approaches to phenomena such as music should examine the relationship between the rules of music and the laws of perception and cognition (p.19). Whereas the nature of sociological thinking is to question if any such universal laws of perception and cognition actually exist; that is laws that can be said to govern all human thinking. This basic conflict between the sociological and psychological perspectives can however be somewhat reconciled. Martin (1995) suggests that sociologists and psychologists are in fact addressing very different questions. He suggests that some psychologists are concerned with the mechanisms of perception and cognition through our sensory apparatus and the anatomy of the brain. In contrast, he suggests that the sociologist is concerned with questions about what people think and perceive. Secondly, the 'laws' referred to by Davies are not physical 'laws' in the sense that gravity is a physical law, but social conventions that have simply been agreed.

Therefore the rules and conventions of social life are neither the result of fundamental mind structures, nor of the physical environment.

“They are socially constructed, maintained and from time to time challenged. Thus they are constantly, often imperceptibly, changing...”(p.7)

This view then suggests that it is necessary to have a *Social Psychology* of music where each parent discipline is given an equal role, with neither acquiescing to the other. Sloboda (1985) suggested that we learn the structures that we use to represent music through our everyday social experiences and through formal training. Whilst Vulliamy and Shepherd (1984) claim suggest that:

“ If the significance of music is irrevocably linked to the *patterning of individual minds*, then it must likewise be linked to the fluid, dynamic and abstract patterning of the social world that lies behind the creation and construction of those minds” (p.60)

The accepted western classical traditional rules of harmony, scales and intervals appear to be a natural and normal basis for organising music in the Western world and yet these ‘laws’ are in fact social constructions and conventions that have been agreed upon and become validated over the last few centuries.

Martin again asks the question that if and when an individual is so imbued with the values and the rules of a particular society (or musical culture), where can a clear distinction be made between the individual and society? He further argues that if the distinction between an individual personality and society is so problematic, any explanation of social behaviour that depends solely upon a sociological or a psychological perspective must also be equally problematic.

Historically, compared to the developmental psychology of music, the social psychology of music has been relatively ignored. Having no similar history, early studies were fragmented and set within no clear conceptual framework. A review of the historical research found that the majority of studies carried out prior to 1955 could be classified according to six broad categories namely: eminence and musical taste, occupational studies, music education, medical and therapeutic uses,

musicianship and composition and ethno-musicological studies. (Hargreaves and North, 1997).

Farnsworth (1954, 1969) produced his book 'The Social Psychology of Music' in an attempt to address what he saw as the imbalance that existed between the dominance of research which explored the biological and physical bases of music and the relatively limited quantity (and sometimes quality) of work exploring the social context and influences in which musical experiences take place. The intervening years between the publications of the two books by Hargreaves and North (1997) and Farnsworth (1954), although not utterly void of studies, produced very little research that built upon the start Farnsworth had made in this area.

A number of studies have explored important issues surrounding music and society; these studies have in turn both clearly signposted the need for and also informed current studies in the social psychology of music. Wishart (1977) and Shepherd (1977) explored the values and meanings of music types in relation to social stratification, whilst Crozier and Chapman (1981) explored the way in which society designates those artefacts that are deemed to be works of art and also their ranking. Vulliamy (1977) presented detailed studies of the process whereby new musics gain legitimacy within society and Leppert (1987) portrays how the music, art, literature and culture of a dominant tier of society mirror the morals, laws and politics of that tier of society (p.104).

There have been a number of further influences that have contributed towards the social psychology of music now being considered more closely. Firstly, the influence of the early debates on social context has brought about a growing awareness in psychology and the social sciences that behaviours should be studied within their social context, if they are to be fully understood. Secondly, the work of developmental psychologists such as Piaget can now be said to have reached a stage beyond the original assimilation of the ideas and moved into a stage where a fuller critique and appraisal of the theories is taking place. Thirdly, the development of new technologies and commercial uses of music have changed both the quantity of music in society and the role that music now occupies / fulfils in society. Lastly, new and young disciplines

and perspectives that were regarded as ‘peripheral’ to more mainstream studies some thirty years ago can now be said to have ‘come of age’. Ethnomusicology, feminism, cultural and media studies, as well as social histories of the arts have become more widely accepted, offering valuable and interesting new perspectives. These perspectives have established culture as a social product, and therefore the study of culture and the arts should accordingly be sociologically informed. Wolff (1987) states that:

“ The notion that art – at least great art – transcends the social, the political and the everyday has been under attack for fifteen years or so” (p.1)

Wolff argues that the aesthetic autonomy of art is a product of nineteenth century ideology and social structure. Sociological changes such as the end of dependency on the guilds; emancipation from direct commissions and improved economic situations enabled artists to become more individual. This was followed by the concept of the development of ‘genius’. Hauser (1962) noted how the concept of ‘genius’ developed within aristocratic societies and works of art became the product of an autocratic personality that was seen to transcend tradition, theory and rules. DeNora (1995) argues that “genius and its recognition require social and cultural resources if they are to be cultivated” (p.xiii). In her book *Beethoven and the Construction of Genius* she portrays the social and political structures, which directly generated the image and status of Beethoven as he is viewed today. She concludes; “ It is an existential fact of life that the social institutions, discourses, and disciplines which enable us to live and communicate with each other simultaneously perpetrate symbolic violence: what is facilitating for some may be constraining for others” (p. 191).

Wolff argues that the construction and perception of all art and artists is inseparable from the society and culture in which they are produced. The nineteenth century myth of art transcending reality she argues, has been perpetuated by interested parties such as commercial producers, specialist institutions and academics. She further argues that even when the aesthetic autonomy of other arts became the subject of sociological criticism, music still remained untouched because of its abstract and non-representational characteristics. In 1987, Wolff argued that published work was at last showing that music does not hold a special place within the arts and that it can

successfully be subjected to sociological, as well as psychological critiques. The ever-increasing body of published work by authors such as DeNora continues to confirm this.

Almost thirty years ago, authors such as Blacking (1973) and Small (1977) both outlined how musical interpretation and meaning is dependent upon the society and the culture that produced it. Blacking commented that a weakness of studies in music psychology was that they had mostly been ethnocentric (p.5). Blacking criticised the efforts of psychologists who attempted to study music:

“ Paradoxically, their laudable aim to be context free and objective fails precisely because they minimise the importance of cultural experience in the selection and development of sensory capacities” (p.5)

Blacking argued that the sensory discrimination involved in listening and musical interpretation is also determined and developed within a culture. Individuals brought up beyond or not immediately involved in that culture may be 'deaf' to some of the more individual features in the music of that culture; or may be unable to express those distinctions. Blacking firmly places any musical activity within the strict confines of the culture that creates it. That is, each culture (or sub culture) has an agreed set of rules and principles about how musical sounds should be organised. Thus, two pieces of music; two notes or two styles may in fact be 'the same' when grouped according to one set of rules, but 'different' when grouped around the rules of another culture pursuing a different social aim. This view was strongly endorsed by Small (1977):

" Other cultures make other assumptions and are interested in other aspects of organised sound” (p.8)

The work of Blacking and Small was concerned with how a society may shape and use musical forms, but they also pointed towards the need for a fuller investigation into how social forces shape psychological reactions to music.

Konecni (1979) also pointed out forcefully that music psychologists had failed to take account of the direct social context in which music was experienced. Konecni argued for a three 'field' model made up of the emotional state of the listener; the musical preferences of the listener and the current social environment of the listener. Konecni claimed all these reacted together and contributed towards the listener experience. The model suggested that the listener was always reacting with the social and non-social environment; reacting to the acoustics of the music, the social behaviour of others, the individuals changing emotional state and individual musical preferences.

In more recent studies, the importance of the social context in which musical events are experienced has been highlighted further (Hargreaves and North, 1997). This research has shown that any measure of musical effect is dependent on a whole variety of social factors. These authors have again argued that in order for psychology to more fully explain the effects and the role of music in the lives of individuals and their cultures, it needed to encompass far more than the effects of physical properties of musical sounds and the ways in which individuals perceived and interpreted them. They argued that the social, cultural and interpersonal contexts within which the music was experienced all contribute to the constructed musical meaning. They further suggest that a number of issues have become more prominent in recent years so that the social psychology of music can no longer be overlooked. Firstly, the development of technological processes has meant that all kinds of music are now available and affordable to all sections of the public in a way that has never before been the case. Music is assuming uses and meanings that are now becoming valid areas for research. Second, the study of behaviour and cognition within a social context has become recognised as being far more important within the fields of psychology and the social sciences (Pollard, 1996). A third influence has been the growing amount of interest from industry and commerce into the commercial ramifications of musical use (Areni and Kim, 1993; Milliman, 1982). Lastly, a recognition that a fuller understanding of musical artefacts can be gleaned through an examination of the society and culture within which the music was produced and presented (Leppert and McClary, 1987; Cook, 1990).

Hargreaves (1986) has reviewed the literature and detailed a number of areas in which the musical values, tastes, preferences and listening habits of individuals can be influenced through social processes and mass culture. The first influence can be termed 'social influence'. According to this view, works of art are assigned value by different social groups. Individual taste is influenced by the desire to align with or conform to the values of that group. (Frances, 1967). A second influence on the preferences and tastes of the individual can be through the status, eminence or authority of the artist. The greater the status of the artist within that culture, then the greater the influences on the way the individual perceives of the music. Third, it has been demonstrated that social class can be a powerful influence on musical taste. Studies by Baumann (1960), Schuessler (1948) and Rogers (1957) all suggested that a preference in children for serious classical music over popular music tended to be found in children from higher socio-economic backgrounds. These various social influences on the individual preferences may act independently of each other or act as a combination of influences.

Lastly, musical fashion may be influenced by the inverted -U curve preference-feedback hypothesis. Studies suggested that the greater familiarity an individual has with a piece of music; the more their preference for it will increase. This continues until a critical point is reached, whereupon further hearings produce a corresponding decrease in preference. Le Blanc et al. (1991) have demonstrated that the degree of familiarity required to reach the critical point and the rate of preference increase/decrease can vary as a function of age and musical style.

2.8 Summary

In this section we have reviewed the rationale behind the numerous calls by social psychologists for the social context to be incorporated into explanations of human behaviours and developmental processes. We have reviewed literature that has called for a fuller range of studies into the developmental aspects of social psychology and we have seen how a whole range of social influences may affect the preferences, tastes and listening habits of individuals. Lastly, we have seen how music can be viewed as a social product and be specifically submitted to sociological enquiry. Certainly, after what appears to be an endless line of 'aborted take-offs', it now

appears that the social psychology of music is establishing itself as a recognised sub discipline. A review of the majority of the literature to date would appear to support the definition given by Hargreaves and North (1997) who suggested that a social psychology of music should deal with the social and interpersonal context in which musical meaning is constructed (p.1).

As we saw earlier, Durkin (1996) suggested 'that social psychologists had come perilously close to overlooking developmental issues'(p.47). He argued for the investigation into developmental courses within specific social contexts; that is a developmental social psychology that drew fully on both the sub disciplines incorporated into its title. A brief review of the literature suggests that with a few exceptions, the majority of research into the social psychology of music has involved adults. A further and growing number of studies specifically investigate adolescents and university students. However, the number of studies within the social psychology of music that utilise participants under the age of 11 / 12 years is extremely small. Perhaps the time is now right to note the criticism given by Durkin (1995, 1996) of social psychology as a whole and for a further sub-discipline to emerge; that of the developmental social psychology of music. Durkin sees developmental social psychology as answering one very simple but important question; simply - how do people come to do the things they do?

Part One: Development of style sensitivity

“Anyone of my age and experience who is both a musician and who thinks about music finds himself in a difficult quandary. One side of it consists in attitude ‘so far and no further’. In other words, it consists in clinging to one’s youth as if modernity were one’s own private monopoly. This means resisting at all costs everything which remains inaccessible to one’s own experience”

(Adorno, 1994, p.269)

Chapter Three

Literature review

In his book *The Arts and Human Development*, Gardner (1973b) proposed a model of human development that attempted to combine a number of theories which had previously been regarded as disparate; in particular the theories of Piaget and Freud. The book also attempted to attribute the arts with a far more important role in theories of human developmental psychology than they had previously been afforded. He argued that full participation in the arts was an acceptable “end state” for development in the arts. Therefore, Gardner developed his own developmental model involving three systems: namely, the making system, the perceiving system and the feeling system. In this chapter, a more detailed account of Gardner’s three systems model of artistic development is presented and this is contrasted with the stage development theory of Piaget.

All developmental models have been the subject of criticism and a number of the main critical issues are discussed in a following section by referring to the work of White (1997). We will then review a brief account of the aesthetic model proposed by Abbs (1987) and note a number of ways in which this model may reconcile some of the discrepancies that exist between the philosophy and the psychology of developmentalism. In the last section, a brief review is given of the four main studies that have to date been carried out into the developmental aspects of style sensitivity.

3.1 Gardner’s three systems model.

Gardner suggested that three “systems” existed which could operate in both independent and interactive ways. Gardner termed these three systems *making*, *perceiving* and *feeling* respectively. The products of the making system are acts or actions; the products of the perceiving system are discriminations or distinctions and the products of the feeling system are emotional and affective. Gardner noted that a limitation of his ‘feeling system’ was that of subjectivity. Whilst actions and perceptions are observable, feelings, affects, emotions or aesthetic responses, can be more difficult to identify. Therefore the same criticism of “intellectual mushiness” could apply here. Similarly, some of Gardner’s ideas can be equally criticised for the

lack of accurate terminology used to describe affective reactions. For example, Gardner used the terms affect, emotion and feeling interchangeably.

Gardner proposed the end state of development was full participation in the arts. Individual participation could be as a performer, an audience member, a critic or an artist. Some individuals would excel in only one of these roles whilst others may gain a high level of competence in a number of them. Gardner suggested that each of these four participatory roles could be developed through varying combinations of his three systems.

- The artist would be highly developed in the feeling system, creating a symbolic artefact to represent that feeling through highly discriminatory choices from the perceiving system and then combine these with highly developed skills of making.

- The critic would be highly developed as a perceiver; making high-level discriminations and turning these into an end product by combining prior knowledge. Critics would need to be highly aware of their own feelings and monitor their own bias.

- The audience member would have a highly developed feeling system. The ability to experience feelings invoked by the artwork but with no need to communicate them to others. Integration of factual knowledge and the affective system could increase or decrease the overall emotions.

- The performer would be skilled mostly in a making system. The performer must have the skills to re make the original creation but also to integrate their perception of the audience requirements and their knowledge of the traditions surrounding the original creation.

" The distinctions (between the systems) are of course rough: The most skilled executants of any of the roles are characterised by high development in them all" (p.41)

Piaget had suggested that infants are born with a small number of reflexive schemata (cognitive and mental structures): new experiences are constantly differentiated and assimilated into these schemata which become more separated, more numerous and much more complex. In contrast, Gardner suggested that the young child possessed three separate and distinct systems. He argued that children could contribute to, and

take part in the artistic process and through physical development and experience, the three systems gradually influenced and impinged on each other and therefore become more and more integrated. Ultimately each system involved the other two in varying combinations.

Gardner argues that a reorganisation takes place where the already now integrated three systems are employed within symbolic media. This reorganisation and employment occurs over a prolonged period of time. However, Gardner argues that the child's increasing adeptness in the various roles of participant in the arts does not need to be explained by proposing any new forms of mental mechanism. Development takes place not through major changes in cognitive development but through participation and experiences within the various symbolic media.

A number of concerns have been expressed about Gardner's model. White (1997) expresses a basic concern, which he claims besets all forms of developmentalism; therefore his caution would be equally applicable to the work of Piaget. He argues that developmental models are based on the central assumption that the '*unfolding development*', which exists in the biological and physical realms, also exists in the mental. Gardner acknowledges two poles in the mental realm of the individual. At one pole are, allegedly, genetically given capacities common to human beings, (visual perception, sensitivity to sounds, bodily movements etc.). At the opposite pole of the developmental process is the end state, the unfolded mature form of the capacity.

Two main problems exist for White, one problem at each end of the developmental continuum. The first problem concerns the initial state or *seed*.

“What is characteristic of biological seeds....., is that they have within them the power to unfold into more complex stages of their organic growth given the appropriate environmental conditions” (p.4)

White argues that biological seeds *unfold* given the right environmental conditions (air, water, temperature etc.). However, environmental conditions in the physical realm cannot be equated with social shaping (culture, background, education etc.) in

the mental realm. Innate capacities in the mental realm, and White does accept that they exist, do not simply unfold as they do in the biological realm, they are socially shaped through the interaction between genetic and social factors. White claims that innate capacities in the mental realm *differentiate*, they do not *unfold*. Therefore whilst it is safe to conclude that a certain plant developed from the innate potential of a particular seed; it is not necessarily the case that an identifiable mental end state developed from an identifiable innate capacity.

White's second problem is with the opposite pole, the mature or end state of development. In the physical realm, a fully grown (developed) end state is an acceptable notion. Any physically developing organism will reach a natural *ceiling*; a limit beyond which it will develop no further. Following this point, the organism may change in numerous ways due to deterioration or maintenance, but it can no longer be said to be developing. However, transferring this notion of development directly across to the mental realm carries with it the implication that humans have intelligence levels beyond which they cannot proceed. That is, mental development has a ceiling White argues that the development of intelligence has no ceiling or limits but " simply potentially endless growth in certain directions" (p.4)

White accepts that this notion of endless development is unheard of in every other developmental realm and that such a phenomena also requires empirical confirmation. However, it does introduce the idea of development moving towards "states of relative maturity" which replaces the idea of a final end states. White challenges Gardner to define clearly what counts as maturity in the case of intelligences. The critical questions for White are firstly that if a mental ceiling exists, how do we know that people have reached it and secondly, if that ceiling does not exist, how do we know that people are more mentally mature than they were.

White contends that judgements surrounding people's intelligence tend to be controversial and also value judgements; that is, judgements that depend upon what the individual thinks are valuable and moral. He suggests that " judgers apply their intellectual and moral values to their decidings – and are likely to differ amongst

themselves because of the different weights they each apply to the multiple criteria that operate in these areas” (p.5).

Throughout the paper, White continues to question the philosophical foundations upon which Gardner places his account of the development of innate mental capacities. He even concludes by suggesting that had Gardner applied philosophical critique to his theories of symbol-theory and to developmentalism, “the original theory probably could not have even reached the drawing board” (p.10).

The underlying problem for White appears not to be the academic debate between two differing perspectives, but the way in which educators appear to take up Gardner’s theories and apply them with great vigour. The United States has a growing number of magnet schools based solely on Gardner’s ideas and theories. White likens this ‘take up’ to the adoption by the teaching profession of Hirst’s *Forms of Knowledge*’ (1972) some twenty years earlier. In this article, Hirst proposed seven areas of understanding, which should be included within a liberal education and White suggests two reasons why these two theories by (Hirst, 1972 and Gardner (1973b)) have been so enthusiastically adopted by educators. Firstly, educators have a keenness for their pupils to experience a small number of forms of thinking / understanding / or intelligence that a leading academic authority has identified as crucial to understanding human nature. Secondly, he suggests that the two theories have struck a chord because they have both tended to be forms of thought that are closer to the intellectual and cultured life normally associated with the educational professions. They are closer to what Lawton (1980) and Bantock (1968) have termed ‘high culture’. I think that White’s main contention with Gardner’s theory is one of epistemology, White is simply questioning the true rationale behind the apparent success of a proposed psychological theory that has not yet been subjected to any degree of philosophical analysis.

In the introduction to the second edition of his book, ‘Frames of Mind’ (1993), Gardner appears to have made a number of interesting shifts in opinion. One main shift, White describes others, is in the acknowledgement of the weakness inherent in

any theory that does not take account of the differences within which human beings live and develop. Gardner also comments on his own clarification of the distinctions between the terms *intelligence, domains and fields*. Two of the three, *domains and fields* are essentially social in construction. This shift is in-keeping with the current focus of situated cognition. Hargreaves (1999) describes this view as one that suggests:

“... the acquisition of knowledge can only be explained by reference to its physical and social context” (p.26)

Hargreaves further notes that this has led to a focus on the interactions between teachers and learners rather than simply concentrating on intellectual changes in learner. Abbs (1987) has taken an educational perspective that stresses interaction, physical and situational factors, and has suggested a model for an *aesthetic field* that he describes as a ‘highly complex web of energy’ (p.55). In this model he identifies four aesthetic areas in which children should partake in order to assist development through engaging with this dynamic aesthetic field. The four distinct areas; those of making, presenting, responding and evaluating have some overlap with Gardner’s realms and roles but Abbs identifies them as follows:

-Making. In this, Abbs takes his definition of art creation from Stravinsky. Making is the impulse to express. This animates the specific medium of the art maker. In the encounter between the two, the art work takes shapes. In the initial stages of the making process, artists are concerned only with their own personal appetite to express and the arranging and modelling of the specific medium within which they have expertise. However, as the work progresses, the maker begins to acknowledge a sense of the audience which adds to the final shape. There is also a readiness to present the finished work to the audience, which gives the work an independent existence. This role corresponds with the role of maker in Gardner’s model.

- Presenting. Here Abbs quotes John Dewey who argued that the finished product was not the work of art. The work of art only exists when other human beings cooperate with it. The presentation, as it applies to music, may take different forms, but the re-creation of the score demands “ fresh and exacting acts of creative indwelling and expressive projection which closely parallel the making process already outlined” (p.59). Certainly in some musical forms such as jazz, the maker and

presenter become one. In short the audience and the art maker require each other. This role corresponds with the role of performer in Gardner's model.

-Responding. Here Abbs argues for a separation of responses into pure aesthetic, that is feelings or emotions without knowledge; and those based more on knowledge. He argues for the use of questions such as "what does it do for you" rather than "what does it mean" or "what is it saying" which lead away from aesthetic responses of feeling into more discursive and knowledge based realms. Responses to art works should be pre verbal, physical and sensuous. This role corresponds to the role of audience member in Gardner's model.

-Evaluating. Abbs calls for a clear distinction between affective responses and more cognitive ones. He accepts that it is desirable to want to judge and understand art works. He also accepts the need for knowledge (historical, technical, factual and cultural) in order to make sense of art works. It is this need for understanding and evaluation that requires the role of evaluating. Evaluating is the attempt to organise the complex elements of our response; the combining of our emotional response with our factual and cultural knowledge. Abbs argues that simply acting the role of critic can put distance between the art work and the responder whereas in the best evaluation, further engagement with the work of art is encouraged in an attempt to further explore and make sense of the sensuous. This role corresponds with the role of critic in Gardner's model.

Initially, it appears that the Abbs' model and Gardner's model are both making an identical claim for four similar component parts. I think the main difference between the two models is in emphasis. Gardner states that as development proceeds, his component roles become more integrated. Abbs however, stresses far more the constant and vibrant movement between the four roles and specifies the way in which the various roles become inter-connected, although in fairness this was beyond the scope of what Gardner was suggesting. Abbs uses the term *field* as a term borrowed from quantum mechanics. It suggests a constant movement across and in between his four areas.

For Abbs, the aim of aesthetic education is to assist development through engaging with this dynamic aesthetic field. Each individual part can be developed through factual knowledge (e.g. through technical or cultural knowledge) but each part of the field gains its meaning through its connections with the other parts. Abbs argues that factual knowledge and experience are constantly turned back into the cyclical process of the aesthetic field. This creates a constantly increasing number of differentiations, but at the same time allows for, recognises and draws upon innate capacities.

Abbs' model appears to support Gardner's model and yet at the same time side step the criticisms made by White. Abbs does this in two main ways. Firstly, he presents his model as educational for aiding development and not as a psychological developmental model. Secondly, the four areas (roles) are to be engaged in for their intrinsic developmental and integrative value regardless of their origins (innate or learned) or their destinations (ceilings).

Gardner (1972a) viewed style sensitivity as a type of concept formation involving perception, developing sensitivity, discernment and the application of rules. He argued that there was one main difficulty in selection of tasks that accurately reflected this kind of human cognition. For Gardner, research tasks that explored concept formation were either controllable and easy to measure but artificial and experimental or they were natural but imprecise, hard to control and hard to define. Gardner (1973a) suggested that little was known about children's perceptions of music or their developing abilities to detect artistic styles. His suggestion was that exploring the developing ability in children to detect musical style was a means of exploring this variety of concept formation whilst at the same time using stimuli that were both natural and yet easy to control in an experimental setting. A review of this work by Gardner and the three studies of style sensitivity that have followed is presented in the next section.

3.2 Gardner (1973a)

Howard Gardner carried out the initial pioneering research on the perceived ability of children to be sensitive to style within the arts. This research (1970) focused on the

way in which children looked at and classified works of art. The results of this research had led Gardner to note that children who were at or below the level of concrete operations had a tendency to focus on either the dominant figure or the participant matter in a work of art. Therefore classification judgements were made according to these features. Further findings by Gardner (1971) suggested that although 7-year-old children had great difficulty in breaking away from this approach to works of art, they could be taught to sort paintings according to style, even if competing figural cues were present. This supported work by Gardner and Gardner (1970) where it had been found that children of 6 years of age could possess a certain degree of style sensitivity to literary and painting styles.

Gardner (1973a) then carried out further research in order to establish whether or not:

" similar trends are discernible in a child's relationship to musical works" (p. 68)

Gardner (1973a) operationalised style sensitivity as the ability to recognise whether two contrasting musical pairs came from the same or from a different piece of music. That is, " the ability to group together works produced by one artist" (p. 326). The musical examples used by Gardner were selected from the standard western classical music repertoire written between 1680 and 1960. His reasoning behind selecting this musical genre was that 1) participants would be familiar with the overall style of the music but would not know the individual pieces; 2) pertinent aspects of the music could be evenly matched and 3) an investigation would be possible of the relationship between the level of style sensitivity and the time gap between the composition of the two pieces of music.

Gardner argued that style sensitivity involved a complex cognitive process:

" which demands monitoring of numerous aspects of a stimulus, avoidance of over emphasis upon a single facet, and attention to the general expressive features of the work" (p.76)

He further argued that the discrimination and classification skills required for participants to be sensitive to musical style may reflect the existence of a deeper and

more central process that was operating within a wide variety of classification tasks that were regularly entered into throughout the individual's daily life. Gardner argued that the central process of concept formation, involved the individual having the ability to detect very subtle differences within a plethora of highly varied data; the ability to then organise it according to complex rules and thus make decisions about grouping objects and ideas together. For example, the simple task of categorising a particular animal as being a 'dog' involves observing identifying and filtering out pertinent data which could identify 'dog'; subjecting this data to an already established data bank of rules and examples which identified 'dog' and then judging if or not the data and the rules were a match. One purpose of the Gardner study was therefore to establish if the way in which musical style sensitivity develops followed a similar process.

There were three further aims to the study. Firstly to establish if participants improved their test performance when given a second hearing of the musical selections. Second to establish if participants could more accurately establish a difference in musical styles coming from widely separate eras and third; to further test the Piagetian idea that younger children would focus on the dominant figure of the music and ignore the textural elements. Gardner used the solo voice as a dominant figure, against an instrumental background as the ground musical figure.

Gardner's original study on musical style sensitivity (1973a) was carried out on groups of children in five different age groups namely: 6 years, 8 years, 11 years, 14 years and 18/19 years old. Participants were 10 males and 10 females in each of the five age groups and were "overwhelmingly middle class and of high intelligence" (p.68). Sixteen pairs of musical selections were played to the participants from four musical styles, namely Baroque, Classical, Romantic, and Modern. Participants were asked to note if the two contrasting pairs came from the same or from a different piece of music. 8 pairs consisted of 2 excerpts from the same composition. Within each of the 8 pairs, 4 pairs were exclusively instrumental and the remaining 4 pairs were mixed with the soloist being present in only one of the two excerpts. All participants, with the exception of the youngest, were told they would have an opportunity to hear

all the examples a second time and after each hearing the participants were asked to give reasons for their decision.

The results of the study firstly suggested that there was generally a high level of sensitivity across all the age groups tested with even the youngest participants showing a high level of accuracy. Indications were that style sensitivity increased with age from 6 year olds through to 11 year olds. The 11-, 14- and 18- year olds all performed at a similar level, but the 11 year olds appeared to score most accurately. In all age groups females had scored at a significantly higher level than the male participants. Further discussion of the results with reference to the main hypotheses for the study revealed:

- 1) Gardner had suspected that given the opportunity to hear the examples a second time would improve participant accuracy, as participants would have the opportunity to re assess their judgement criteria. This was not proven. There was some evidence that elder participants may have benefited but younger participants did not. The 11 year olds showed no improvement and the 8 year olds scored worse. Gardner suggested that this lack of improvement or deterioration in younger participants might be attributed to boredom and possibly fatigue.
- 2) There was some evidence that participants did identify with greater accuracy those musical styles coming from non-adjacent eras. Within the two older age groups, participants had a significant tendency to identify items from widely divergent eras. The three younger age groups also exhibited this tendency but not to a significant level.
- 3) There was some evidence for the projection that younger participants would gravitate towards the dominant figure as the basis for their judgements. This trend did not exist as strongly as in other studies involving paintings with some participants apparently not noticing the voices. Gardner attributes this to the fact that other factors present in the music (e.g. rhythm and volume) may act just as much as a dominant figure and that an equivalent 'dominant figure' may be less pivotal in music than in painting.

The verbal responses for participant's choices of 'same' or 'different' were also analysed. Gardner found that the verbal reasoning behind participants' judgements increased qualitatively and not quantitatively. The youngest participants could give little or no reason for their responses. Typical responses were of one word describing one aspect of the music. e.g. loud or soft; no further reasoning could be extracted when subjected to further questioning. However, as Gardner was quick to point out, there was no evidence as to whether or not this was due to lack of style perception / discrimination or lack of the verbal skills required to express the reasoning.

The eight year olds expressed their decisions according to criteria beyond music terminology. Adjectives such as " churchy" or "peppy" were used suggesting that participants were linking the musical excerpts with their own personal experiences in other domains. The 11 year olds showed awareness of a number of different variables and were the first age group to acknowledge that two discontinuous excerpts could in fact come from the same composition. This age group appeared to listen to the music in its own terms, describing what they heard rather than trying to attribute it to a historical period or assign aspects of it to musical terminology.

The eldest participants tended to show a developed awareness of musical style and described the reasons for their choices and classifications with appropriate music / technical language. This age group attempted to assign aspects of the stimulus to musical categories and terminologies rather than just describe the stimulus itself. Gardner cited instances where more sophisticated responses gave rise to inaccurate responses. In a typical instance, a participant would correctly identify that two pieces came from the same era but then vacillate as to whether the pieces were from different movements of the same piece or were actually by two different composers within that same style. They could often identify correctly a possible composer for the piece. It was sophistication of this level that Gardner (1973a) felt might be a source of error in the older participants.

He suggested:

"This excessive introspectiveness may be one reason why, using an alternative approach, less sophisticated eleven year olds did not perform significantly worse than undergraduates" (p.74)

Hargreaves (1982) also noted that children's tendency to classify music in terms of stylistic labels increases with age. These findings, although set within a different task context give additional support to Gardner's results.

To sum up, Gardner argued that young children do have some capacity to make stylistic judgements with some tendency to centre their judgements on certain dominant figures. He argued that two approaches could be effective in musical style detection, the first one for younger participants being their responses according to personal experience and attention to the actual properties of the musical stimulus. The second approach for older participants being their attention to musical categories and musical terminology.

Limitations

Gardner himself made a number of criticisms of his research. Firstly, he noted limitations in his research population. The study used a relatively small number of participants, (ten males and ten females in each of the five age groups). All were taken from one social class and ethnic background and selected in the hope that they would be " favourably disposed towards classical music and motivated to perform well on a difficult task" (p.68). All participants were labelled as "highly intelligent". Gardner suggested that the middle class participants used in the research population may produce results that were not generalisable to a wider population. Second, the research incorporated only classical musical excerpts. Therefore no comparison was possible either in the participant's potential performance in other styles or with the results from other research. Third, the participants were offered a reward.

3.3 Castell (1983)

Castell (1983) followed up Gardner's' study by making a number of additions and modifications. She set out to explore to what extent children were sensitive to musical styles and attempted to identify those features of the music, which children noted in their process of discrimination. The research followed the same basic rationale and procedure as Gardner's study but made a number of refinements some of which had been suggested directly or indirectly by Gardner himself in the discussion of his research results.

First, the musical examples were changed. Gardner had noticed and commented on the ambiguity of the dominant musical 'figure' and this was therefore changed. The dominant figure as expressed in the Gardner examples had been thought of as the prominent feature of the piece. i.e., the melody played by solo instrument. Gardner had however suggested that the labelling of the dominant figure was essentially 'adult' oriented. Whereas it appeared that participants had in fact identified their own individual dominant 'figure' as that within the music, which was most familiar to them. This was evidenced by the fact that in some instances younger children even failed to notice the presence of the voice in the example, and based their decision on other noted features. It was argued by Castell that the issue of how a 'figure' relates and interacts with a musical 'ground' was a complex debate within itself and perhaps this was complicating the test in an unnecessary way. Therefore, no vocal line was used in the Castell examples. All examples were instrumental music excerpts with an equal number of instruments involved in each piece.

The second modification made by Castell, was the introduction of an equal number of jazz and popular musical pairs which were included along with the classical examples. The rationale for this was two-fold. Popular styles were felt to be pertinent in this study because they would be more significant to the age groups involved. Secondly, it was noted that popular musical styles had very much been under-represented in previously published research. The general attitude of researchers towards popular music was summed up by Wing (1968). In his tests of musical ability and aptitude, Wing had argued that the inclusion of popular styles in his tests would be 'unlikely to yield examples of really good harmony'; that it would 'prejudice the authorities against the tests' and also 'waste the children's time if they were listening to poor music' (p.37).

Therefore both popular and classical musical styles were included in equal quantities in Castell's research. Twelve pairs of musical samples were used (6 classical pairs and 6 popular pairs). The recording criteria, i.e. the duration of playing time and separation times for the pairs followed that used in the Gardner research. An additional feature of the research was that the test was set within a story in order to make the test more user friendly. Gardner had closely followed the test introduction

style as used by Pflederer Zimmerman (1970). This was formal, short and gave a simple factual account of what would take place within the test, i.e. that the participant would hear a piece of music followed by a second piece. Participants were then simply asked if they felt the two pieces were the same or different pieces of music.

Castell changed the test introduction by situating the instructions within a story setting. Participants were asked to imagine that they were able to open a door to a room and hear the music within played by musicians. Participants were then told to imagine that they had returned to the room some time later. They were then asked if the musicians were playing the same piece of music or if they had started to play a new different piece. This change in introduction addressed the possibility that younger participants could in fact be experiencing some 'linguistic difficulties' connected with their understanding of the key words 'same' and 'different' or even misunderstanding the instructions totally. The use of the story added a visual and concrete element to the instructions and removed the verbal emphasis of previous research.

In addition to recording whether or not they felt the two excerpts were from the same or a different piece of music, participants were also asked if they felt the piece could in fact have been written by the same composer. Further to this, participants were asked to note briefly the reason for their choice. Castell's research also included participants from a broader social and economic spectrum than Gardner's initial research, which had used only intelligent, middle class participants. Also, there was no offer of reward.

The results of Castell's study suggested that 9-year-old children had scored more accurately than the 11-year-old participants. Machotka (1966) had posited that style sensitivity would only begin to appear at around the age of 11 years old with the child's developing ability to conserve. The fact that Castell had produced results contrary to this prediction prompted her to carry out a second study which extended the age group by using participants aged 7 to 18 years plus. Castell suggested that factors other than age may have affected the results of the first study. Namely that:

-The style sensitivity test was measuring an ability different from the "style sensitivity" as defined by Machotka or that style sensitivity does actually decrease with age. Therefore in order to explore this issue further, the age range was increased to include younger participants (age 7) and older participants (over 18 but a mean age of 21.2yrs.). and a few minor alterations were made to make the test more age friendly with respect to the older age group. Other factors were kept the same.

The results of Castell's two studies provided an interesting contrast to the Gardner study. The results suggested the following:

In study number one which explored style sensitivity in children aged 8 to 11 years:

- As with Gardner's results, the overall standard of performance by participants was high.

- The main effect for age showed that the 8/9-year-old participants performed better than the 10/11-year-old participants.

- Both age groups performed better on the popular pairs than the classical pairs. Further analysis showed some interesting interactions between age and musical style.

- Both age groups performed equally well on the classical pairs. The increased performance of the younger children was due to their making fewer errors on the popular pairs.

- In contrast to Gardner's results, participants performed differently on type of pair. Gardner had discovered that his older participants reduced their overall scores by a tendency to over sensitivity. That is, they appeared to note pairs came from a different piece when in fact they came from the same piece. There was an over sensitivity towards any small change in the piece of music. The study by Castell found the opposite to be true in that 11 year old participants appeared to note that excerpts from the popular pairs were from the same piece when in fact they came from different pieces.

The fact that 9-year-old participants made fewer errors on the popular pairs than the 11 year olds, was a result that deserved some further comment. Chapman and Williams (1976) and Murdoch and Phelps (1972) argued that peer group norms may be an important factor amongst adolescents in determining reactions to music. Type of music would appear to be an important group identifier amongst adolescents who exhibit a tendency to be very knowledgeable about narrow styles of music and the artists themselves. May (1985) gained a measure of participants' preferences for popular, non-western and art music. His results suggested that adolescent preference ratings decreased with age across all styles, with the result for popular music showing the least pronounced decrease across all styles. particularly for Rock and Country music. Greer, Dorow and Randall (1974) presented results that suggested that preference for popular style music increased gradually between nursery and 12 and 13 year old children with what they described as a 'critical' change in preferences between 8 and 9 year -olds. They also demonstrated a parallel decrease in preference for classical, symphonic and orchestral music. LeBlanc, Sims, Siivola and Obert (1996) produced similar results. Children aged from 5 years to 13 years showed a decrease in preference for classical, jazz and rock music. Similarly, Hargreaves, Comber and Coley (1995) suggested a decrease in preference ratings for popular and classical music as age increased. However, preference ratings for popular styles showed the least pronounced decrease. Brittin (2000) explored children's preferences for different types of sequenced electronic keyboard accompaniments to a number of well-known children's songs. In this research she also found a " trend of diminishing 'open-earedness' through the elementary grades.....with the older students giving the lower preference ratings." (p.246)

So, Cástell (1982) reported that the poorer performance by the older participants only occurred in the popular pairs. In an attempt to explain this age effect, she therefore suggested that the older participants might well have been affected by very clear boundaries in listener tolerance.

“ By the age of 11 children have formed very definite likes and dislikes in the field of current pop music, and cease to distinguish very much between pieces and styles that fall outside their chosen favourites” (p.25)

and those younger participants might well be said to be:

"...more 'open-eared' as there might be less social pressure on them to like certain types of music and dismiss others" (p.25)

In one sense, it seems that the expanded term originally used by Kreitler and Kreitler (1972) of "a more open – eared and open-hearted listener" (p.144), is slightly more descriptive of the effect. This decrease in open-earedness is supported by Rodriguez and Webster (1997) who suggest that their own research found that children demonstrated a gradual devaluing of the opinions of their music teachers with adolescents strongly favouring peer group norms. Older participants developed very clear and often narrow ideas of their tastes in popular music, these tastes being defined by their own peer culture. Furthermore, it was suggested by Castell that the tendency of older children to be 'over inclusive' may be influenced by the fact that older children are more aware of the wide range of varying musical possibilities existing within one single composition or style.

This issue raised by Castell requires further comment and clarification. The main point being made by Castell is that all her participants performed better with popular styles with which they were more familiar. She further suggests that the poorer performance by the 11-year-old participants in the popular pairs might be due to a decline in 'open earedness'. It therefore needs to be asked by what process did this age effect influence the participants' performance. There are two possible reasons. Firstly, the decline in 'open earedness' may have caused participants to be less tolerant of those styles beyond their immediate preferences and they therefore listened less attentively. That is, participants dismissed those styles beyond their immediate 'likes' and did not fully attend to the musical excerpts. Secondly, the reduced level of accuracy might have been due to the tendency of the older participants to be 'over inclusive'. That is, they were too sensitive and reacted to even minor differences in the musical excerpts and therefore judged the two excerpts to be different when they were in fact the same. There is also the possibility that both effects were present during the full duration of the test.

If the first suggestion were to be a major influence, it might be a reasonable expectation that this lack of participant tolerance – and therefore attendance- to those excerpts beyond their immediate ‘likes’, might also occur in some form of musical style effect, as well as age effect. That is, participants might also be expected to hold the same lack of tolerance for, and poor attendance to some of the classical excerpts, which might also be well beyond their immediate ‘likes’; thus creating both an age and musical style effect. Similarly, if the second suggestion were true and 11 year-old participants were ‘over sensitive’, it is reasonable to expect that this would also reveal itself amongst the classical excerpts, simply because popular music tends to be more uniform throughout than classical music. Therefore, in the popular pairs there should be fewer differences to be over - sensitive to.

In reality, it is probably very difficult to separate age effects from stylistic tolerance / musical style effect. It is probable that the age differences could be a result of stylistic tolerance changes with age, and / or participants’ differential familiarity with the styles. In terms of the experiments carried out in this thesis, the idea that participants’ musical tolerances change according to individual musical styles is accepted as the hypothesis and it is this idea that is tested and explored further. At a later point in this thesis, an alternative explanation for the difference between participants’ scores in the popular and classical pairs is explored.

In the second study by Castell, the age group was extended to include children aged 7 and 18 years old. The main results of the study suggested that:

- style sensitivity did increase with age. Adult participants did perform more accurately than other participants and 7-year-old participants did make the most errors.

- 8/9-year-old participants did perform better than 10/11-year-old participants. This result may be a reflection of the fact that a random sample of the total number tested in the first study.

- There were more correct responses for popular pairs than for classical in younger participants with adults scoring equally on both types of pair. It was therefore

suggested that the ability to tell whether two excerpts came from the same composition was affected by style of music until a certain age (somewhere between 11 and 18 years old.)

- Adults and 7-year-old participants both showed a tendency to be over inclusive, saying pieces came from the same piece when in fact they came from different pieces. The age groups in between showed the opposite tendency.

Castell therefore concluded that style sensitivity did increase with age with participants appearing to be more familiar with popular styles rather than classical styles. Adults gave the most accurate responses with the 7 year olds making the most errors. The mean score of the 9 year olds was higher than the mean for 11 year olds and also participants scored more accurately on popular than classical pairs except within the adult age group where responses were equal in both classical and popular pairs. In spite of the differences in mean scores, she argued that children firstly acquired the concept of style at around 9 years old with participants showing a high level of accuracy in matching musical pairs. With 11-year-old participants, Castell agreed with Gardner that their concept of style was becoming too broad. Participants were becoming less tolerant of the styles they were prepared to listen to their increasing knowledge allowed for the possibility of too much variation within a single piece. By adulthood, the responses had become more accurate. Sadly, as there were no participants between the ages of 11 years and 18 plus years tested in this research, no conclusions were possible on the development that takes place between the ages of 11 and 18 years.

So the results suggested by Castell present an interesting comparison with the Gardner results. To sum up, the two Castell studies suggested that style sensitivity increases between the ages of 7 and adulthood with younger participants giving more accurate responses to popular music. Also the tendency to give correct responses to 'same' pairs increases with age, whereas the tendency to give correct responses to 'different' pairs decreases with age.

Limitations

A number of limitations existed in Castell's study, which could affect the generalisability of the results:

- a random sample of the 9 and 11-year-old responses from the first study were used in the second. No new 9 or 11-year-old participants were tested. If these results from the first study were somehow atypical due to extraneous reasons, this atypical group of results was transferred to the second experiment.

- although not an aim of the study, the fact that no participants were tested between the ages of 11 years and 18 years old, meant that no data existed within this important area of development. Hypotheses about the nature of the development of style sensitivity between these age groups were therefore limited.

3.4 Tafuri, Addessi, Luzzi and Baroni (1994, 1995, 1996)

Tafuri, Addessi, Luzzi and Baroni undertook and reported on a number of their own studies (Tafuri, et al. 1994; Addessi, Luzzi & Tafuri, 1995). In these two studies, the authors concentrated on classical musical pairs and also explored a change in test format using three musical samples instead of musical pairs. Three of the same authors also produced a fuller and much broader report (Addessi, Luzzi & Tafuri, 1996) which incorporated both the change in test format and the use of both classical and popular pairs. The final report (1996) incorporates some of the pertinent features of the two previous reports plus some additional material. In total, the three reports present a somewhat confusing account of the research with individual points not always being followed through from one account to the next. A summary is attempted here.

The study incorporated some similarities to the previous studies by Gardner (1973a) and Castell (1983), but also incorporated a number of differences. The first part of this study again operationalised style sensitivity as the ability to identify correctly whether two pieces of music came from the same or from different musical pieces (Test 1). The second part of the study incorporated an alternative and possibly more robust method of operationalisation. In this format, (Test 2) participants were asked to judge which

one of the three musical fragments was not written by the same composer. The expectation from incorporating this technique was that the presence of a more stylistically homogenous pair should present a much clearer background against which to judge the odd one out. This should produce fewer incorrect 'different' responses and therefore a corresponding increase in overall accuracy scores. The study incorporated the same four classical repertoire styles that Gardner (1973a) had used and in the later study, much broader modern and popular styles were included. Different classical and popular pairs were used for the pairwise and triadic tasks.

One further difference between the studies by Tafuri et al. (1994), Addessi et al. (1995, 1996) studies and the previous studies by Gardner (1973a) and Castell (1983), was the somewhat different definition applied to style sensitivity. The studies by Tafuri et al. opted to incorporate the term *stylistic competence* and not *stylistic sensitivity*. Their definition of stylistic competency is somewhat complex and draws on the work of Nattiez (1990). A brief report on their use of the terminology is given here.

Addessi et al. (1995) argued that concepts of musical style were complex and unclear. They argued that there are features that distinguish one style from another i.e. classical from romantic or baroque; and there are features which are common between very different individuals within one style, i.e. Mozart and Haydn. In that sense, musical styles can be considered as being both 'the same' and 'different'.

In the process of creating, all artists make choices. They work within a culturally agreed set of rules and regulations and select the degree to which they abide by or depart from those agreed rules. However these choices are made for expressive reasons which reflect the emotional life and the ideology of the artist. When a number of consistent choices are repeatedly made, this may be termed as the 'identity' of the composer. Therefore;

“ Artists are able to create a ‘musical style’ when they are able, with their creativity, to manipulate the traits of the already existent grammar (in a given artistic tradition) to allow for the symbols of personal identity. Thus, we can say that style is *the symbolic representation of identity*” (p.9)

From the perspective of the listener, to perceive style is not simply to recognise its structures. Addressi et al. (1995) argue that to perceive style is to interpret style. That is, listeners need to perceive specific features of the music but must add to this perception an internal operation before they can speak of it, that is: interpret it. 'Interpretation' is giving sense to the original perception and this depends upon the listener having specific forms of knowledge. The more syntactic knowledge the listener possesses, then the richer will be the interpretation of the music. The identity traits of the composer, as recognised by the listener are referred to as *stylistic pertinences*.

Therefore, for Addressi et al. (1995) musical stylistic competence in the listener can be defined as:

“ the ability to identify stylistic pertinences in a piece of music, that is, the ability to recognise an identity through musical traits that symbolically represent it” (p.10)

Stylistic competence is a more complex process than style sensitivity, which has been operationalised as the ability to identify whether or not two excerpts came from the same or a different piece of music.

The stated aims of the studies carried out by the Italians vary between the different reports, but a summary is presented here.

- To verify the development of musical stylistic competence in children.
- To address what is meant by stylistic competence and how does this develop.
- What aspects influence this development and can it be promoted by music education.
- To explore if stylistic competency increases proportionally with participants' knowledge of a particular music repertoire.

They further hypothesised that a greater *stylistic competency* would be required to identify music from periods, which were chronologically closer together. That is, the

closer together in time that two representative pieces of music were written, then the greater the competency required to identify the different elements.

Participants for the study were drawn from 6 age groups namely 8,9,10,11,12 and 13 years of age. Test 1 consisted of 15 pairs of musical selections by the same or different composers. Test 2 incorporated 18 groups of three musical selections. In each of these groups two selections were written by the same composer whilst one selection was written but a different composer. The 'Test 1' (Pairwise task) classical musical selections were taken from various concerti for solo violin and orchestra whilst the popular excerpts included rock (Bon Jovi, Iggy Pop, Pete Frampton); Disco music (Plastika, Flower Factory, Outhere Brothers); International music (Prince, Yazz, Big Country, Witney Houston) and Italian song (Clio, Pupo, Zarillo).

In 'Test 2' (Triadic task) the classical material all involved pieces written for solo piano and taken from the same four classical periods namely; Baroque, Classical, Romantic and Modern. The popular excerpts were all taken from Italian popular song; the songs were all taken from one of four decades, namely: the 60's, 70's, 80's and 90's. (Lucio Battisti, Edoardo Bennato, Fabio Concato and Ligabue). These selections were chosen in order to present excerpts that were more homogenous in instrumentation with the excerpts differing in style only. In order to examine the hypothesis that it would require a greater stylistic competency to identify musical selections from periods that were chronologically closer than those periods which were chronologically more distant, all excerpts in both the pairwise and triadic tasks were assigned to a difficulty category (*livelli*). That is, those musical selections from adjacent eras were judged as being most difficult whilst those musical selections from distant eras were judged as being easy.

The popular pairs were placed in a rank order. This order was not established in any empirical or controlled way. It was simply decided upon by the authors own ideas as to how stylistically different they considered the various excerpts to be; "the categories were ranked in approximate order of stylistic distance" [*Le categorie sono state qui elencate in un approssimativo ordine di "distanza stilistica"...*] (p.63) . The

Italian songs were judged, by the writers, as being closer to the international music excerpts with the rock music being ranked next – equivalent to two eras away from Italian song. The disco music was placed last in the category as this was felt to be the most stylistically different from the others. Therefore, a contrasting pair containing Italian song and Disco, were judged to be a similar in difficulty (*livelli*) to Baroque and Modern, i.e. three eras away.

Results

In common with previous studies, the results in this research again suggested a high level of accuracy amongst all the participants with even the youngest participants scoring significantly higher than could be expected through chance. The results of the various studies are reported in slightly different ways. The study reported in 1994 included results on classical excerpts only. The results suggested that:

- The main effect for age was not significant.
- Performance was better in styles that were familiar in both the pairwise and triadic tasks.
- Stylistic competence did not vary according to stylistic differences resulting from distance in time, but did appear to vary as a result of stylistic divergence from those stylistic identities with which participants were most familiar.

The 1995 study suggested different patterns between the results obtained in the pairwise and triadic tasks. In the pairwise tasks there was a difference between age groups but no significant increase due to age. In the triadic tasks, the difference in performance between age groups and the increase in performance due to age were both found to be significant. The results did lead them to consider that stylistic competence did increase according to familiarity because participants did perform better on those excerpts they were most familiar with; that is Baroque music and excerpts by Mozart. The authors claimed greater participant familiarity with these musical styles through their prominent use in advertising, media and film sound tracks.

Accordingly, the results from these studies suggest that stylistic competence is dependent more upon musical acculturation than on the age of the listener. Although the role of music education is stated as a question, no comment on the role it may play in the development of stylistic competence appears in any of the three studies. Leaving the change in test format aside, the results seem to suggest that participants perform more accurately on musical excerpts which are more familiar. That is those styles for which they have an already established internal pattern. This would appear to have major implications for music education in that stylistic competence should improve if participants experience more and greater varieties of music.

In the classical pairs / triads, participant's performance did not vary as a function of stylistic divergence between styles resulting from chronology, but they did appear to vary as a function of stylistic divergence from those styles with which participants were most familiar. In this case Baroque and Mozart. The authors argued that subjects had 'internalised the patternings' of Baroque and Mozart music because it feature more in the media and was therefore more familiar. The real question to be asked here however is this; did participants acquire the internal patternings of Baroque music and Mozart because they heard the music more frequently in the media, and it was therefore more familiar or could it be said that some intrinsic feature of Baroque music and Mozart causes these patternings to internalised first?

In the later study (1996), more accurate responses were evident in the popular excerpts than in the classical excerpts on both the pairwise and the triadic test formats (63.9% -72% classical – popular pairs and 41.4% - 57 % on the classical – popular triadic format). It was in this last report (1996), that the writers compared the classical and popular pair results for the first time, so they added a further comment on the effect of the levels of difficulty resulting from whether styles were adjacent to each other in time.

“it was more difficult to recognise different styles between neighbouring eras; but this was not confirmed with the popular material, where more than the style of an era counted as a difference between genres which were naturally more familiar”

("... e piu difficile riconoscere differenze stilistiche tra brani di epoche vicine, ma cio non vale con il materiale popular dove piu che lo stile d'epoca conta la differenza tra I generi, oltre naturalmente alla familiarity" (p.79)

That is, the adjacent / non-adjacent effect was not present in the popular pairs where listeners may have been aware of other musical features within the musical excerpts. It may also reflect the fact that the writers did not group their popular material in any systematic way.

Limitations

The three reported studies by the Italian authors were excellent in two main ways. Firstly, they introduced a new and interesting comparative methodology, in the form of the pairwise and triadic tasks. Secondly, they introduced the interesting and alternative terminology of 'stylistic competence' to further describe the mental process taking place within the style sensitivity test. However, there are a number of limitations to the reports that may be mentioned.

One possible limitation is the large amount of musical material that the study incorporated into both the pairwise and the triadic tasks. Whereas previous studies had used a small amount of musical material to represent a musical style, this study used very different pieces by different composers within each style. Sometimes up to four different composers were used to represent a single musical style, and in a number of musical aspects each of the composers within that style is, or can be classed as being as different from each other as they are from composers of other styles e.g. Prokofiev and Berg or Beethoven and Haydn. As the number and type of cues that participants use to make their decisions about style sensitivity judgements is still unknown, the introduction of such a large amount of material is a possible source of confusion.

Similarly, the study incorporated a considerable amount of music by Prokofiev as a representative of modern music. Musically speaking, Prokofiev himself frequently protested about the severe limitations placed on his work by the Russian government. Certainly, the work of Prokofiev could be placed in a number of dominant stylistic categories. In terms of the hierarchies of style put forward by Meyer (1989);

Prokofiev may be classed as being a strong member of seven categories namely: Works by a group of composers; works written in the same geographical location (drawing heavily on Slavic roots); works written in a particular genre; works written for some socially defined segment of a culture (to the glory of Russia); works written for some utilitarian purpose (Russian ideology); works assigned to a recognised period (modern western) as well as the works of the individual composer himself. As such, and in terms of the question asked of participants in this experiment, he can be classed as being 'the same' as Tchaikovsky and Stravinsky depending on the musical elements that the listener attends to.

It was the authors' stated aim to establish:

" The ability to identify stylistic pertinences in a piece of music, that is, the ability *to recognise an identity* through the musical traits that symbolically represent it."

(Addessi et al. 1995, p.10)

The authors are attempting to establish how and when children are able to identify and recognise musical patternings, that is the identity within the music. Accordingly, in the 'Test 1' material, musical samples are employed that are all " extracted from "Allegro" of different Concerti for violin and orchestra in order to avoid the influence of other variables (agogic,dynamic etc.) on the results" (Tafari et al. 1994, p.308). Secondly, the authors argue that a number of levels of 'difficulty' may be assigned to the various musical pairs. That is, it should be easier to identify differences between musical styles that are from non-adjacent periods because of a greater stylistic divergence. Similarly they argued that participants had performed better on musical styles with which they had an already existing internalised pattern with which to compare the musical samples (e.g. the Mozart pattern obtained from the media). However, it is argued here that the ability to identify and establish 'patternings' is subject to other levels of difficulty than simply noting the level of contrast brought about through chronological distance between the musical eras. It is far easier, for example to identify and internalise the replicated patternings of Baroque and Mozart musical samples, than to accomplish this same level of identity in composers operating within the same, yet more modern eras. For example, the authors used the music of Berg, a composer whose serial musical patternings / identity are extremely difficult to identify, even by an informed and competent musician. Often the

‘patternings’ of this style of music is eventually isolated and ‘explained’ by a mathematician rather than a musician, let alone an ordinary listener.

It is therefore questioned here as to whether or not young participants of six years-old, in spite making a correct response, have the ability “*to recognise an identity through the musical traits that symbolically represent it*”; at least in the case of contrasting musical examples by Berg and Prokofiev.

3.5 Hargreaves and North (1999)

Hargreaves and North (1999) carried out a further study into style sensitivity. In this study the authors used the same classical musical excerpts as in the previous research by Addessi et al. (1996) but not the same popular excerpts. Participants were taken from one primary and one secondary school in the East Midlands and were aged between 8 and 14 years old. The research was described as a "further test of the effects of genre" by using the more robust triadic technique, as used by the Italian study. A new batch of popular examples was added. These excerpts were taken from mainstream Blues, Grunge, Thrash and Indie and represented popular genres which participants would find familiar. The testing procedure was similar to those used in the other studies.

The classical excerpts within each style were all taken from works by the same composer and all included solo piano. All excerpts were kept as homogenous as possible with respect to instrumentation, volume and tempo so that stylistic variation was the only salient difference. Correct detection by certain specific combinations of triads could therefore be ruled out. Each style appeared as the 'odd one out' once. The popular examples were pieces taken from the works of a single artist. All featured electric guitar and did not feature voice. The classical triads were always presented first in random order, followed by the twelve popular triads, again in random order. The same random order was kept the same in each test situation. Participants were presented with a response sheet and told they would hear 12 trios of musical excerpts. In each case one would be the odd one out.

The study revealed some further interesting results. The first hypothesis was not confirmed as there was no clear effect for the interaction of age. This finding was in line with the results of Addessi et al., but did not conform to the results of Gardner and Castell. It was suggested that comparison between the results of the three studies may be difficult due to the variance in age range, genre and testing method in the three studies.

This prompted Hargreaves and North (1999) to suggest:

" Any further tests for age effects need simultaneously to incorporate familiar and unfamiliar genres; pairwise and triadic techniques, and a sample drawn from the full age range of these studies" (p.199)

The second hypothesis of the research was supported with participants gaining higher scores on the popular music examples, possibly showing higher levels of stylistic sensitivity on genres that were probably more familiar to and liked by the participants. This effect very strongly outweighed any effects due to age in this particular study.

In the Hargreaves and North (1999) study, there was strong evidence that the triadic task was discriminating effectively, as evidenced by the levels of scoring. However, the Italian study by Addessi et al. had noted a possible decrease in the apparent accuracy in the triadic testing method as against the pairing testing method. The study by Hargreaves and North consisted of only the triadic task and therefore no further comparison of levels of discrimination between the pairwise and triadic method was possible. Similarly, the results reported in the present thesis consisted solely of pairwise tasks, so again no further comparison is possible between the two methodologies. Again, as Hargreaves and North stated; any further tests "need simultaneously to incorporate familiar and unfamiliar genres; pairwise and triadic techniques" (p.199).

3.6 Summary

The four studies by Gardner (1973a), Castell (1982), Addessi et al. (1995) and Hargreaves and North (1999) provide a mixed set of results. The absence of any clear

cognitive advances coupled with the greater accuracy in the popular genres as opposed to the classical genres lead Hargreaves and North to speculate that:

" social and cultural factors exert a powerful influence on any potential cognitive determinants of style sensitivity.....it is probably impossible to disentangle the cognitive aspects of responses to musical styles from their social and affective components" (p.200)

The present research detailed here does not claim to disentangle the cognitive aspects from their social and affective components, but the following experiments do attempt to investigate further a number of those components and to un-pick some of the social and cognitive aspects by further investigating children's musical stylistic sensitivity in a number of different contexts.

Chapter Four

Methodological considerations

“If the personal characteristics of the data collector have determined in part the subjects’ responses, then we must hold our knowledge the more lightly for it. There are experiments by the dozen, which show that different experimenters obtain from their comparable subjects significantly different responses”

(Rosenthal, 1976, p.39)

This chapter is divided into two main sections. The first section (4.1) reviews some of the previous literature which has explored possible interactions between test participants, the test presenter and the testing environment; whilst the second section (4.2) reports on the test procedure and the selection of musical material used in experiments one to four of this thesis.

4.1.1 Research precedents

As Rosenthal (1976) has pointed out, in any test situation there is always a high risk that some form of uncontrolled interaction will take place between the test presenter, the testing environment and the participant. There is therefore a concern that this interaction will significantly affect the final results. A substantial research base exists which catalogues studies that have explored some of the effects of these interactions that may be present in any test situation and ultimately may affect the motivational level and the emotional state of the participant experiencing the test. In this chapter a number of these previous studies are reviewed however, no claim is made that this review is a full and comprehensive account of all the literature available (see Anastasi, 1988; Sattler and Theye, 1967; Cronbach, 1990). Rosenthal (1976) identifies two kinds of experimenter effect. Firstly, he identifies experimenter attributes that are independent of the experiment. This category would include biosocial attributes such as race, age and sex of examiner. Secondly, there are those experimenter characteristics that can only be defined in terms of the actual experiment. This category would contain attributes such as status of experimenter to participants, warmth of experimenter-participant relationship. Further to this he highlights a difference between those effects that are passive and those that are active. Active effects occur due to unintended differences in experimenter behaviours that can be

seen to influence participants' responses whilst passive effects are those where no such behavioural differences are present and therefore may be due to biosocial attributes alone.

In practice these different styles and forms of influences are difficult to isolate and invariably there is always going to be a combination of the two. The literature suggests that there are a great many variables that affect the participant's response other than those variables, which, in any given experiment, are specifically under investigation. The personality of the experimenter, how he or she looks and acts, may by itself affect the participants' response. The effect can be direct and simple but other times it can interact with participant characteristics, task characteristics, or situational characteristics.

In the present study a range of experimenter attributes were considered during the planning stages of the experiment design. As a result of this some modifications were made to the overall test design in an attempt to limit, account for or control any resulting influence these attributes might have had on the experiment results. Following Rosenthal's categories, both passive / active and dependent / independent attributes were considered. These attributes are reported in the following sections. The second part of each section details how the overall methodology was changed in an attempt to limit each particular effect. The studies and attributes reported in this chapter are of a general nature and refer to all experiments carried out in the present study. A number of further experimenter attributes that are more specific to individual experiments are reported in following, more appropriate, sections.

4.1.2 Experimenter rapport

Exner (1966) reported that experimenter rapport was a significant influence in testing situations. Improved participant performance in the form of higher overall scores in intelligence tests were reported as a result of the behaviour of the experimenter prior to the test itself. In Exner's study, two conditions were operated. In the first condition the experimenter established a degree of rapport with the participants prior to presenting the test. This involved the presentation of a warm personality, e.g. smiling, showing respect for participants and entering into some form of non-test related

conversation prior to the test. The second condition involved the same experimenter operating the test but presenting a cold and clinical personality with neutral expressions and a neutral attitude to participants. Verbal contact was limited to test instructions presented in a formal way. The study reported significantly higher scores from participants experiencing the warm personality than from those participants experiencing the cold and clinical personality.

Hata, Tsudzuki, Kuze and Emi (1958), had previously carried out similar investigations of this effect. These authors carried out a similar study to Exner but had grouped participants according to different personality traits. Their study again used intelligence tests and the results suggested that the effect of pre test rapport could vary according to the personality of the individual participants with not all personality types equally affected. Cronbach (1990) noted that experimenter rapport was far more important in situations where individuals were being tested rather than in a group test situation. The testing of individual participants is pertinent to later experiments in this thesis and is discussed in more detail in a later section. Additionally, Cronbach suggested that within any test situation, the relationship between experimenter and participants could be much improved by incorporating praise, general approval and smiling. All three of these suggestions were incorporated into the present testing format.

There is some suggestion that experimenter relationship is important. "Familiarity, understanding, warmth, preference, and adjustment all play a role in altering the participant's performance" (Sattler and Theye, 1967. p.355). Throughout the experiments in the present study, each experimenter aimed to establish a relationship with the participants by use of a number of pre-established test neutral topics. These included comments about how much the presenter always enjoyed listening to music as a child; how nice it was to be back in such a pleasant school and some comment about the weather. All presenters then engaged in a short dialogue with the participants, each asking a number of short questions. The questions involved a showing of hands as a response to the question. The questions asked how many people had a radio; if they listened to it every day and how many had heard three particular stations. (Virgin radio, Kiss FM and Classical FM). This pre-test question



session was introduced as a result of the pilot study. The pilot studies had noted that younger children had a tendency to volunteer information, for example, "My brother has a radio like that" or "my dad has got that music". Although this was taken as an indication that some rapport had been established with the participants, it was inappropriate for these comments to come during the instructions. As all three presenters had each witnessed the other two presenting the test, it was felt that a reasonable level of agreement over the level of 'rapport' existed. It was also felt that the use of a more 'participant friendly' introduction to the test; that is by incorporating the instructions within a form of story about the radio, created a beneficial pre test rapport.

4.1.3 Experimenter expectations

Dyer (1973) reported differences in test performance as a result of the difference between the experimenter and the participant's perceptions of the main aims of and the functions of the test itself. Results by Masling (1965) reported differences in participants' test scores caused by differences in the experimenters' expectations of the participants. Typically this has been where the experimenter has been given pre test data on the participants. For example, the experimenter is given false test results that are artificially high. The experimenter therefore expects greater performance from participants' scores. Further studies (for full review see Lutey and Copeland, 1982; Anastasi, 1988) suggest no conclusive evidence but it was noted by Lutey and Copeland that:

" there is enough evidence that such effects do exist to be alert to the importance of this factor" (p.133)

In view of the nature of the test in the present study, it was felt that it was appropriate to acknowledge this effect.

A significant finding of Dyer's research was that follow up discussion between participants and test presenters showed that neither had any awareness of how they had either influenced or been influenced. Subsequent tape recordings of the introductions given by the various test presenters revealed that there were no identifiable differences in the language used, nor any identifiable diversion from the set procedure. It was therefore concluded that the experimenter's expectations

apparently operated through subtle postural and facial cues to which the participants responded.

Rosenthal and Jacobson (1968) carried out a major study on the effects of teacher expectations on students' performance. Their findings suggested that personal interactions were somehow affected by expectations that in turn could lead to self – fulfilling prophecies in performance. They felt able to speculate only on how this was achieved; possibly through teachers being more friendly, pleasant and encouraging with children from whom they expected greater intellectual gains. They suggested that greater attentiveness and closer supervision might have led to more rapid reinforcement of correct responses. This reflective behaviour on the part of the teacher may in turn have been learned by the pupils, which brought about a change in their cognitive styles. They concluded that by:

“ What she said, by how and when she said it, by her facial expressions, postures, and perhaps by her touch, the teacher may have communicated to the children in the experimental group that she expected improved intellectual performance” (p.180)

Throughout the experiments described in this thesis, some attempt was made to control expectations. Experiment two involved the use of three different test presenters. The purpose of the experiment was to examine the possible influence of presenter characteristics on participant's tolerance of different musical styles. It was therefore felt to be desirable that some effort be made to control experimenter expectations as an influence on participant performance. Therefore the following steps were taken. All three presenters were equally aware of the justifications for the experiments and the functions of the test. All the three presenters witnessed each other operating the test during the pilot study, and all three presenters had some individual input into the final test procedure. Following the pilot study, all three presenters took part in marking the response sheets and some discussion took place regarding the level of accuracy and performance of the participants. Each marked some of the response sheets and the responses and further expectations of the test were discussed. No academic data on participants was seen by the presenters and all instances where classes had to be divided into equal ability groups, the class teacher or Headteacher did the division.

4.1.4 Pre test experiences and activities

Research by McCarthy (1944) and Reichenberg-Hackett (1953), reported a significant influence on intelligence test performance in relation to pre test activities.

Participants' scores were significantly lower in tests that followed a pre test activity designed to provide a negative and emotionally depressing experience. Likewise, Davis (1969) observed a poorer performance in participants who experienced intelligence tests following a pre test activity designed to make participants experience failure. No corresponding increase in performance was found when participants received a similar amount of praise and perceived success prior to the testing situation. This 'failure' effect was found to be greater on younger children (Lantz, 1945).

Konecni (1982), reported on pre test experiences and their reported influence specifically on listener tolerance. Konecni carried out an experiment where an accomplice of the examiner throughout a pre test activity repeatedly insulted participants. Subsequently, these participants were given a choice of listening to complex or simpler melodies. Participants who had been insulted displayed the tendency to select the less complex musical samples to listen to rather than the more complex ones. Konecni accounted for this effect by suggesting that participants in an already high state of arousal, as caused through the insults, opted for less complex music. By making this choice the participants opted not to increase their arousal further through listening to the more complex musical examples. That is the listener tolerance to some musical examples was affected by prior activity.

Whilst success or failure in an activity immediately prior to the testing situation can affect participant performance, some researchers suggest that issues of success and failure may be part of a learned behaviour pattern that has been firmly established over time. O'Neill and Sloboda (1997) investigated children's affective and behavioural responses with specific reference to music testing. In particular, their investigation focussed on the effects of test-failure on pupils performance in a test of melodic direction. In one trial, pupils encountered success, whilst in their second trial pupils encountered experimenter induced failure. The results suggested that pupils who had

experienced failure also experienced a deterioration in their performance on subsequent tests, whilst those experiencing success either improved or remained the same.

In the present study, no control was possible as to whether participants experienced success or failure prior to their taking the style sensitivity test. Neither was any control possible over the emotional experience or arousal level of the participants prior to their taking the test nor the music teaching patterns that participants were accustomed to. However, some consideration was given to the manner in which the practice test examples were handled. This is reported in greater detail later.

4.1.5 Experimenter status and experience

Davis (1969) found that status of experimenter did not have a significant effect on participants' performance in intelligence tests. However, examiner experience can be a source of influence. A number of studies have explored the possible effect of training and experience of the presenter on participants' test performance. (Sattler and Ryan 1973; Grossman, 1978). In these studies, and in others reviewed by Lutey and Copeland, no significant effects were found in the performance of participants with an experienced or inexperienced test presenter. The one realm in which experience did appear to affect results was in the area of test marking with experienced and inexperienced testers scoring greater accuracy on marking different forms of test material. Rosenthal (1976) suggests:

“ it seems reasonable to suppose that a more experienced experimenter, one who has conducted more experiments, or at least repeated a certain experiment more often, may behave differently in the experiment than a less experienced experimenter” (p.90)

Rosenthal reports on research that has investigated the effect of experimenter experience on participant performance in a task of evaluating expression of success and failure in photographs of faces. The main findings suggested that more experienced experimenters present the verbal instructions for a test in a less animated way than less experienced experimenters. Rosenthal suggests that this change occurs through boredom. Interestingly, although there is some change in behaviour between the more and less experienced experimenters, there was no effect noted on the

participant's ratings of the photographs. Furthermore, he suggests that the experimenter may change behaviour during the course of a single experiment.

As Rosenthal (1976) concludes the amount of and type of experimenter experience may influence test results, these studies were noted and a number of changes incorporated into the design of the experiments reported in this thesis. Firstly, as experimenter presentation appeared to change as a result of the number of times the experimenter had presented the test (possibly through boredom), the particular experiment involving a number of different experimenters, was placed early on in the testing schedule. In this case therefore, all three experimenters had a similar amount of experience in administering the style sensitivity test. Secondly, selecting the other experimenters, as in experiment two, individuals were selected who were confident characters, had previous experience of the school environment and children and were comfortable with a musical stimulus. Third, the previous research was shared with the other experimenters so they were aware of possible behaviour changes and lastly, all three experimenters observed each other present the test in a pilot study. In terms of status, all schools were asked not to 'announce' the visit during the morning of the test or prior to the visit and all experimenters were introduced as being of equal status, i.e as music teachers.

4.1.6 Test anxiety

Anastasi (1988) reviewed research on the effects of anxiety on participants' performance in test situations in tests. She concluded that a number of procedures, if included within a test situation, might contribute towards a lowering of anxiety levels. Procedures that dispel surprise and strangeness from the testing situation; assurance and encouragement; good organisation and communication and experimenters manner may, they suggest, all contribute towards a situation of reduced anxiety. Mandler and Sarason (1952) found that tests involving ego-boosting instructions had a beneficial effect on low anxiety participants but a negative effect on high anxiety participants.

In order to minimise test anxiety in the present experiments, care was taken with the terminology of the introduction, the procedure and the printed material. At no time was the word 'test', or any variant of this, used. Presenters were always introduced as

'Music Teachers' who had 'come to do some music' with the class or group. Participants were informed that responses were not to be marked as for 'right' or 'wrong', but would be interesting in terms of what they had heard 'in their own opinion'. All presenters were experienced with children and presented a warm and relaxed attitude to the participants. (Exner, 1966). Schroeder and Kleinsasser (1972) have pointed out that where participants are unknown to the test presenter then experimenter effects may be reduced. This may be due to the fact that the outcomes and therefore the ramifications of the participants' performance are more distant and therefore may be less anxiety provoking. In most cases, the teacher remained in the classroom and therefore the participants were not simply left with a stranger.

4.1.7 Experimenter distortion

Cronbach (1990) has highlighted the important fact that any test situation is a social relationship (p.72). As participants will be susceptible in different ways to experimenter effects, so experimenters will also be subject to testing effects. One major concern posed here by Cronbach is that of post-test distortion through tester bias. Cronbach points out that bias in interpreting resulting data is especially likely when certain scores, " will bring benefit or blame to the person collecting the data" (p. 78). With respect to the present studies recorded here, there was not felt to be any significant anxiety over whatever forms the final results took. The data from previous research had presented a number of discrepancies so there was not felt to be any pressure to necessarily conform.

4.1.8 Summary

There are numerous studies that have attempted systematically to vary test conditions in order to investigate possible experimenter and environmental effects on participant performance in test situations. A number of those most pertinent studies have been detailed here. Broad reviews of the whole literature base such as those by Satler and Theye (1967) have often pointed to results which are contradictory, non significant or contrasting. Sattler and Theye (1967) explain these often-contradictory results by pointing to the large number of methodological inconsistencies which exist between studies.

" The major difficulty in evaluating the research findings with respect to experimenter variables is that experimental procedures were frequently less than adequate" (p.354)

This view is also endorsed by Anastasi (1988) who comments that although many of the experimenter and variable effects reached some level of significance, many others were inconclusive or misleading due to the failure of the experimental designs adequately to control the individual variables. These authors have suggested a number of factors that could have contributed towards the numerous discrepancies. These are certainly applicable to the studies quoted here.

-Many of the previous studies have involved a number of variables that were not held consistent between the various methodologies. For example, different styles and types of testing materials were employed by the individual studies. Reviews of the literature suggest that test materials may easily differ in their susceptibility to experimental procedures. Authors such as Schwartz (1966) observed that levels of susceptibility towards experimental effects could vary within a single test according to the ambiguity of the question. Schwartz reported tests involving ambiguous questions appeared to be most susceptible to experimental effects. So the effect of test, test environment and experimenter influence on participant performance may actually vary within one single test.

-There was a difference among the previous research studies in how and when the deformations to the standard were administered. Some treatments being administered prior to the test; some during the test and some throughout the test.

-Many of the studies draw on research carried out in individual testing situations. Very few of the studies carried out further studies to assess if the results were equally applicable to testing in-group situations.

-Almost all previous studies had used some form of intelligence test. Effects may therefore not directly transfer to testing situations requiring responses that were less 'logical', more creative and more open ended.

Therefore, if much of the research examining experimenter and test environment effects has resulted in such contradictory findings, in what sense can this literature be

relied upon? Several points should be raised here. First, the majority of previous studies have incorporated testing materials designed to assess intelligence, verbal or mathematical skills. There is insufficient evidence to suggest that tests involving listening to 'musical samples' will necessarily have the same susceptibilities to experimenter and test environment effects. Secondly, discrepancy of results is not in itself cause to assume that a particular effect is insignificant. Identification of weakness in previous methodology can create more resilient methodologies in the future. Total separation from what has gone before because of discrepancy of results, may in fact change the entire paradigm. Like is not compared with like. Lastly, further research will constantly produce new data, which may create new ways of interpreting previously recorded data.

To summarise, there is a wide diversity of influences that may affect a participant's performance in any test condition. Throughout this study, every effort was made to be aware of and acknowledge the probable sources of external influence, which may affect participant performance. Adequate and reasonable steps were taken to minimise their influence. Even a brief review of the literature in this area highlights a wide diversity of variables, which may, to varying degrees, affect individuals in any testing situation. There may be significant interactions between experimenter and participant characteristics. Cronbach (1990) has pointed out that it is insufficient to simply isolate individual variables, which may affect participant performance. The same experimenter and experiment may produce different effects on different participants as a result of the participant's own characteristics or even the participants emotions and mood on the day.

Anastasi (1988) suggests that:

" In the majority of well administered testing programs, the influence of these variables is negligible for practical purposes" (p.40)

Regardless of this suggestion, the concluding advice followed in this present study has been for the experimenter to be:

" constantly on guard to detect the possible operation of such variables and to minimise their influence" (p.40)

4.2 The present studies: general methodology and procedure

In this second section, a description is given of the test design and methodology used in experiments one to four of this present thesis. There is also a detailed section as to how the test musical content was selected and recorded. The section also contains details of the way in which the response sheets were marked.

4.2.1 Test design and modifications

Within the design of the test methodology for experiments one to four, an attempt was made to address some of the limitations that were apparent in some of the previous research as outlined by Hargreaves and North (1999). They had suggested that further tests should involve similar age groups to those used in the previous research. Similar age groups were therefore found and used in experiments one to four although no 18-year-old plus participants could be found in sufficient quantity.

Gardner (1973a) used predominantly middle class, educated participants. In common with the Castell's (1982) study, the present studies incorporate a wide range of participants from different cultural, ethnic and social backgrounds. A variety of schools at all levels were sought which differed in size, catchment area, geographic location and with varying degrees of musical activity.

Castell (1982, 1983) had reported a possible source of distortion to the results of her second study. In this study the participant age range was extended. In two of the age groups, the data was achieved through random withdrawal of a number of participants' responses from her first study. Castell noted that any distortion present in these particular results in the first study could have been carried on into the second study. Castell further noted that some difference had occurred in the testing environment of two of the age groups; one group had been tested in class-sized groups whilst the second had been tested in smaller withdrawal groups. She concluded that some participants might have found one test environment more 'friendly' than the other. In the experiments detailed here, no research population or result was used twice. This was done in order to increase the sample size of the research population. As a result of this approach, it was felt that any atypical results caused by extraneous factors at the time of the testing or other influences pertinent to one individual school

could be highlighted more easily. In each experiment a new control group was used and both conditions were carried out within a short time period of each other. Some schools were used twice. Where this occurred, subsequent experiments were carried out during a different academic year.

Musical excerpts were taken from familiar and unfamiliar genres but no attempt was made to examine the relationship between the results gained through pairwise and triadic testing. It was felt at this time that this feature would introduce a new and complicating variable.

4.2.2 Reliability and validity

Any measure needs to be reliable. In the case of a research instrument, reliability requires the instrument to produce consistent result. That is, the same results or measures should be obtained when participants are retested at some later point, or when a different researcher applies the instrument of measurement to a similar population.

In addition to consistency, any research instrument should show validity, that is the instrument must be seen to be measuring exactly that which it is designed to measure. Researchers generally signify two forms of validity; namely internal and external validity. Internal validity questions whether or not the experimental treatments do in fact measure that which they were designed to measure, whilst external validity questions whether the measure relates predictably to external criteria of what is supposed to be measured.

Cohen and Manion (1980) have detailed a number of threats to external and internal validity. Each of these 'threats' is examined here with respect to the research proposed in this study.

Internal Validity.

1. History.

This effect relates to events which happen over time between the pre-test / initial test and the post-test / subsequent test. In terms of the present research there was a time lapse of between six and twenty four months between the initial and the later data collection. Musical trends and fashions can often be short lived and therefore participants in the later cohort could have certainly been influenced by a very different musical / cultural diet. In the case of the youngest subjects, up to 25% of their total life experiences may have been substantially different. Therefore, it is possible that their view of, tolerance of and perception of the same musical examples as used in the initial pilot and early cohorts, may have changed considerably. It is also the case that the first four experiments reported here, in an attempt to control music as a possible variable, used the same examples as those used by Tafuri et al. (1994). Essentially this places the participants in the latter populations some 5 or 6 years apart from the original Italian populations. Considerable changes in teenage culture, fashion and musical trends would most certainly have taken place and this fact could most certainly threaten the internal validity of the instrument of measurement proposed here. One point to consider with this must be that in replicating the experiment some 3 years later, Hargreaves and North (2001) used the same classical samples of music and yet found some reliability amongst the two sets of results.

2. Maturation.

This threat relates to the fact that the same subjects may change due to maturational processes between two specific points of observation. It may therefore be these maturational changes that affect resulting data in a pre-test/post-test situation and not the actual variables. This was not considered to be an issue in the present study as the participants were not observed over a prolonged period of time.

3. Statistical regression.

This threat is that in pre-test and post-test situations, there is often a regression to the mean. Although there is no use of pre-test / post-test within this particular experimental design, there is still a possibility that regression to the mean could impact upon the findings.

4. Testing.

Pre-tests or practice examples can produce practice effects which result in higher scores on the subsequent test scores. Alternatively, they may produce ‘sensitivity’ effects to the test itself. (Rosenthal, 1976, Bernard, 2000). That is, practice examples may present the participants with alternative information, or clue as to how the test response is to be processed, or they may engender a particular attitude towards the content or test methodology which influences the participant’s test performance. As the test methodology used in experiments one to four of this thesis only included two brief practice questions, this was not felt to be a serious threat from practice effect. However, it is possible that the particular styles / examples of music used in the practice examples, coupled with the extensive language content of the introduction, could have produced a sensitivity effect which did influence the participant’s tolerance to the test in general.

5. Instrumentation.

Unreliable tests can introduce serious errors into experiments. It has to be accepted that the measuring instrument used in the first six experiments of this thesis may lack internal validity. Essentially, all previous studies carried out in this area of research have used a similar instrument of measurement and therefore may have been subject to identical spurious elements within the test instrument itself. This is an important issue and indeed one of the main aims of the experiments 1 to 6 is to question and explore further the validity of this form of test instrument when used in this way. Writers such as Castell (1982) and Tafuri et al. (1994) made some substantial claims based on data resulting from experiments using the form of test instrument used here. Experiments one to six of this thesis attempt to further explore the validity of this test instrument by attempting to systematically examine a number of the variables which may in fact be influencing participants’ responses beyond artefactual elements of the musical style.

6. Experimental Mortality.

This issue more frequently affects longitudinal studies and refers to drop out rate of participants from the original sample. Participant drop out can ultimately affect the final results by biasing the sample. That is, those participants who remain with the study are in some way significantly different from those participants who drop out of

it. In terms of the present study, the number of participants who opted out in a real physical sense was not an issue. All participants took the test within a school lesson and all completed the test process. However, the test may have been subject to a degree of mental opting out throughout the duration of the test itself. That is by referring to the definitions given earlier, some participants may not or may cease to listen to the musical examples and begin to, or only hear the musical examples. That is those participants who are willing and able to listen to all samples in all eight styles throughout the test may represent a biased sample. Mental opting out through lack of tolerance of the test itself or of the test material may have therefore threatened the validity of the test.

External Validity

Cohen and Manion (1980) present the following possible areas of threat to external validity.

1. Failure to describe independent variable explicitly.

The number of independent variables present within the test situation may be far greater than those described as being under observation in each of the experiments. There are certainly a number of elements in each of the testing situations which may have exerted an influence on the participants' performance and yet have not been identified. Factors such as mood, personal preferences, time of day and attitude towards test presenter may all influence performance, but are not identified and therefore accounted for. Secondly, in a number of the experiments some of the independent variables may not have been explicitly described. For example, in the second experiment, three different presenters of the test were used and the resulting data appeared to suggest that some of the effects could be gender related.

There were significant differences between the three presenters in addition to age and gender. Other significant differences between the presenters could have produced significant influences in the test situation. There were certainly personality differences between the two younger presenters; one could be described as quieter and more considered, whilst the other could be described as buoyant, overt and very confident.

Similarly, the older presenter had some 20 years of teaching experience with 8 years as a headteacher, whilst the two younger presenters had limited experience of schools.

Likewise, although three contrasting schools from contrasting regions were selected for the USA and the UK, it is not claimed that any of these schools were representative of the large variation to be amongst educational establishments in either country. Therefore, it is accepted that the external validity of the testing procedure could have been threatened in this way.

Although more variables were present within each of the experiments, the argument here is that no previous study has made an attempt to explore those variables which may in some way affect and influence the test performance of participants in tests of style sensitivity as operationalised here. Whilst it would have been possible to further explore the influence of the test presenter, for example, by increasing the number of experimental conditions to include other categories such as older- female, teacher – non teacher, extravert- introvert, this was not done in this instance due to the already large number of research populations required for the full study. The aim of the experiment was therefore not to make any substantial claim for the effects of gender or age of presenter, but simply to explore whether or not the test instrument was sufficiently sensitive to this effect and therefore to have influenced participants' performance in previous experiments. If this proved to be a possibility, then it would be worthy of further studies which incorporated a greater number of experimental conditions. This issue is explored further in the final chapter, where all the results are reviewed

2. Representativeness of sample.

Although an experiment may use a truly representative sample of the research population, the population from which the sample is taken may not be fully representational of the greater population and this may affect the generalisability of the resulting data. Certainly, whilst it can be claimed that a large research population was used in the study reported here, several issues may be raised within this category of threat. Firstly, whilst a large number of schools were approached to take part in the research, a relatively small number of schools actually did agree to the request. As a result of this, some schools were used to supply two or more research populations

throughout the study. It is certainly not possible therefore to suggest that the resulting data can be generalised to other schools even within the immediate geographical location. Those schools who agreed to take part may have done so because of a particular interest in music and musical activities. This could easily have been reflected within the ethos of the school and the context within which participants therefore viewed musical activities.

3. Inadequate operationalisation of dependent variables.

Dependent variables must have validity in the non-experimental setting to which the findings are to be generalised. That is, although the dependent variable may be adequately defined and interpreted in the mind of the researcher, participants may not view the actual test activity as valid or necessary beyond the test itself. This is certainly a valid criticism of the present test activity because there is no evidence that participants view, describe, judge or indeed categorise musical styles outside of the test in the way they are required to do so within the test.

Construct Validity.

Construct validity concerns the extent to which the method of testing actually tap into the concept under investigation. The dependent variable under observation in the study reported here can be summarised as the “ability to classify music”. Construct validity is improved by taking a variety of measures of the same concept. As all previous studies had incorporated an almost identical measure of the concept of style sensitivity, construct validity may have been quite low. Certainly, one major criticism of this form of style sensitivity test is that there is often an assumption made that all judgements made within the test are arrived at through the same mental process. For example, in the Italian studies, the authors argue that a number of levels of ‘difficulty’ may be assigned to the various musical pairs. That is, it should be easier to identify differences between musical styles that are from non-adjacent periods because of a greater stylistic divergence. Similarly they argued that participants had performed better on musical styles with which they had an already existing internalised pattern with which to compare the musical samples (e.g. the Mozart pattern obtained from the media). However, it is argued here that the ability to identify and establish ‘patternings’ is subject to other levels of difficulty, rather than simply the level of contrast brought about through chronological distance between the musical eras. It is

far easier, for example to identify and internalise the replicated patternings of Baroque and Mozart musical samples, than to accomplish this same level of identity in composers operating within the same, yet more modern eras. For example, the authors used the music of Berg, a composer whose serial musical patternings / identity are extremely difficult to identify, even by an informed and competent musician.

It is therefore questioned here as to whether or not young participants of six years-old, in spite making a correct response, have the ability “*to recognise an identity* through the musical traits that symbolically represent it”; at least in the case of contrasting musical examples by Berg and Prokofiev.

4.2.3 Pilot studies

Two pilot studies were run in two of the age groups namely: age 11 and age 14. These were carried out prior to the actual experiments. The pilot studies were carried out in schools that were not used in the subsequent experiments. The introduction, answer sheet, equipment and conditions were as close as possible to those used on the actual experiments. Neither the school nor the participants were aware that the research was a pilot study. The purpose of the first pilot study was a) to familiarise the presenter with the equipment, the introduction to the test and the materials; b) to test the equipment for clarity of reproduction and comfort of volume within a classroom environment and c) to establish that discrimination was taking place within the musical examples chosen.

The second pilot study allowed the two alternative presenters, as used in experiment two, to become familiar with the process, the equipment and the introduction. Both alternative presenters were music students at a local college. Both were training to become teachers and both were familiar with and comfortable with visiting a school and a classroom. Each of the two presenters observed the other administering the test in order to align their presentations as closely as possible to each other. There were no further differences between the two pilot studies. In each pilot study, the answer sheets were marked and checked to ensure that discrimination was taking place. In the first pilot study, one musical sample, that of classical against classical had been inaccurately perceived by all participants. This was therefore changed before the

second pilot study. During the second pilot study, approximately half the participants correctly perceived this comparative pair and therefore this was used in the actual experiments.

Following each pilot study, 4 participants in each age group were asked to comment on the test procedure. They were questioned specifically about the introduction. Asking the participants to explain the task as if they were taking the test themselves did this. As a result of this, several changes were made to the vocabulary of the introduction. As a final check, the samples were played through in a classroom setting on two different machines and one was selected for greater clarity of all pieces. Although in the majority of cases, the presenter knew the schools and the Headteachers, no participant had any prior knowledge of the research.

4.2.4 Musical materials in experiments one to four

The musical material for the first four experiments used both classical and popular pairs and was selected with reference to the previous studies done by Hargreaves and North (1999) and by Castell (1982). The method for recording was selected according to the criteria set down in Gardner's (1973a) original study and closely followed by other writers. Every effort was made to contact the authors of all the previous studies in order to obtain the original comparative pairs. Although this study was not in itself a replication of the previous studies, it was felt that using the same musical samples might remove one possible variable. Hargreaves and North (1999) had suggested that discrepancies between results of the previous studies might be due to some artefactual element in the musical examples that participants had reacted to. Alternatively, the examples chosen were poor or non - effective examples of the style they had been selected to represent.

However, none of the previous studies could provide the musical examples they had used due to the time period that had elapsed. It had been hoped that by using the same examples as two of the previous studies, one possible area of error could be removed.

4.2.5 Musical content in experiments one to four

The popular examples used were taken from the same source as the popular examples used by Hargreaves and North (1999) but not necessarily in the same combinations. Each sample used only instrumental music featuring an electric guitar. Examples of four styles were used namely: Grunge, Thrash, Blues and Indie. Full details of the excerpts are listed in the Appendix. 15-second samples were recorded with a 15 second silence between the two examples.

For the classical examples, the same criteria for recording was used as had been described by Gardner (1973a) in that no sample was taken within a minute of the start or the finish of the piece and each sample was taken from within the same movement of the piece. The procedure for selecting the starting point of each classical sample differed from the previous studies. For each recording within the classical samples, a computer-generated sequence of random numbers was used. Each classical excerpt was assigned a random number. These numbers dictated the amount of time that elapsed from the start of the piece to the moment that recording began. Any number generated between 1 and 5 was used to signify that recording would commence 60 seconds after the start of the piece. Likewise any number generated between 6 and 10 used a 90 second time period before recording began. Numbers between 11 and 15 used a 120 second period and numbers 16 to 20 used a 150 second time gap before recording took place. By employing the use of random numbers, it was hoped to limit human bias by allowing no opportunity for subconscious selection. Through this method, all musical features were ignored.

In the classical samples, where the same piece of music was being used, the first section was recorded followed by a fifteen second break. The recording then continued after the fifteen-second break, the music being 15 seconds further on. The starting place being dictated by the computer method described above. In the contrasting pairs from different pieces, the first piece was recorded for 15 seconds again with the starting place selected by computer generated number. Following the 15-second break, the second piece was recorded for 15 seconds. Again, the starting point for this was also dictated by computer generated number. Therefore it is possible that the first piece in the contrasting pair offered a sample of music that

occurred over 2 minutes into the piece whilst the example set against it, although coming second in the tape, actually came closer to the start of the overall piece it was taken from.

The samples were all placed onto cassette tape. The same reproduction and recording equipment was used and the same reproduction equipment was used in all four tests. The material was thus recorded; 15 second sample from piece 1; 15-second silence; 15-second sample from piece 2; 10 seconds silence.

The 10 completed classical and popular pairs were presented to participants in random order. Previous researchers such as Castell had also presented their examples in a random order mixing classical and popular pairs together. Hargreaves and North, on the other hand, had presented participants with all the classical pairs followed by all the popular pairs. Hutt (1947) had discovered that by alternating hard and easy items within a test produced a significant increase in some participant's performance scores. On balance therefore, it was decided to present a random order of mixed classical / popular pairs. The random order was decided by assigning each pair with a number. A computer was used to generate the order in which the examples were used.

In the present study, the same methodology was used to select the musical samples as had been utilised in the previous three studies. However, it is accepted that the nature of the randomisation of musical extracts may have led to some extracts from the same piece being more or less similar depending on where the randomisation fell within each of the two pieces. This factor may be particularly evident in the classical pieces which can contain far more variety within a piece than in the popular pieces. This most certainly could have affected decisions about whether or not the extracts were in fact from the 'same' or from 'different pieces' of music. This is an interesting question and is covered further in chapter nine, particularly in section 9.1.3 and in Meyer's levels of analysis. The specific issue regarding whether or not the exact extract from each piece can and does affect a participant's response was taken into account when reviewing the construct validity of the instrument of measurement. A full review, criticism and re-designing of the test design is presented in chapter 10.

4.2.6 Response sheet

An identical response sheet was provided and used by all the first four experiments. All pupils were given one answer sheet on which to record their responses. The sheet requested details as to gender and age, but no pupil was individually identified. The participants were not asked for any form of written responses and no reading of instructions was involved. For each sample, the participant was asked to circle the letter "S" if they thought the examples were from the same piece; the letter "D" if they thought the examples were from a different piece and the letter "Y" if they thought "Yes, I know this piece". In the case of the youngest age group, children were invited to use a ruler in order to keep within the same line as the number of the question being looked at, at that precise moment.

In order to prevent any form of mix up between answer sheets, the sheets were marked with a 'Batch Number' which gave the initials of the school, the name of the class and the letter 'N' or 'V' representing 'Variable' or 'Non' variable activity.

4.2.7 Testing procedures

A similar procedure was used in all four experiments. An identical procedure was used in the first two experiments, as detailed here. In experiments three and four, the only difference involved a slight change in the introduction. Some children were tested in small groups whilst others were tested in a full class situation. When this occurred is detailed in the individual experiment reports. At the start of each session the tester was introduced to the class as a 'Music Teacher'. Where this did not occur, or where there was some change to this method of introduction, this is stated in the individual test. There was no mention of the term 'test' at any time. The teacher moved the children a reasonable distance apart and the response sheets were given out. Pencils and rulers were checked.

Some general comments and interactive questions were then asked of the participants, as has been detailed earlier.

At the start of each test situation the participants were told the following:

Today, we are going to listen to some music. Have you ever listened to a radio and heard music playing like this.

(A short 5 to 10 second excerpt of a current radio station was played once).

Have you noticed that sometimes you like the music and sometimes you do not? Have you noticed that if you turn the tuner on the radio, you can sometimes get the same station at a second place on the dial?

(This was demonstrated. A short 5 to 10 second excerpt was played of the second station once. The radio was positioned so the re tuning could be seen by the participants. Radio 2 was always used, as this was always available in a number of different locations on the frequency band and in each geographical area.)

That is because the same station can appear at more than one place on the dial. Sometimes, you can tune the radio and a different station appears. You can tell this by either different music being played or by a different person talking

(This was demonstrated by tuning to Radio 3 once. A 5 to 10 second excerpt of the new station was played once).

This radio can also record. Last week I did this; I recorded some music from one station on the radio, then I turned the tuner until I found another station. Sometimes I found the same station playing the same music and sometimes I found a different station playing different music. I made a tape of this.

I want you to listen to this tape today. You will hear the first piece of music from the first station and then a period of silence whilst I tune the radio until I find a second station, just like you have seen me do today. Sometimes, the second station is playing the same music and sometimes the second station is playing different music.

When you listen to the tape, if you think that the stations are playing the same music and therefore are the same station, put a circle around the 'S', which stands for 'Same'

If you feel that they are playing different music and it is therefore a different station, put a circle around the letter 'D' which stands for different. If you know one piece of music, put a circle around the letter 'Y' which means, 'Yes, I know this music'.

We will now do two practice answers. Look at the sheet (or put your ruler) under the line which says 'Ex. 1' [Note: for the younger age group this was written on the board and was pointed to]

(The first example was then played and the participants wrote down their answer)

Good, did anyone have any problems with that? Here is the second practice. (Move your ruler down one line and.) Put your answer on the sheet where it says 'EX.2'.

[This was again written on the board for the younger age group]

(The second example was played and participants wrote down their answer)
Well Done! That was the practice examples of music. For example number one; who thought they were from the same piece; good; and who thought they were from a different piece, good; it is good to see you are writing down your thoughts and not just those of other people. Example number two, who thought the same; good and who thought different, excellent.

Now before we start, does anyone have any questions?

The whole procedure took approximately 25 minutes. The papers were then collected.

4.2.8 Marking procedure and analysis in experiments one to four

The same marking procedure was adhered to in all of the first four experiments. In common with the four previous studies, the response sheets were given a provisional inspection. All sheets where the participant had clearly not understood the task, that is where a participant had frequently circled both the ‘same’ and ‘different’ or frequently crossed out both responses were withdrawn. Secondly, in order to ensure an equal number of responses across both conditions and in all age groups, the class teacher withdrew a number of response sheets at random. There were a number of reasons for this random withdrawal. Firstly, the number of male and female participants in each class group was often very unequal. In withdrawing some responses, some effort was made redress this imbalance. Secondly, it had been decided to follow previous studies and to use Tukey HSD tests as a part of the further analysis of participants’ mean scores. Tukey HSD was regarded as a more powerful tool than most other multiple comparison tests and operated most efficiently when comparing cells of equal or almost equal sizes. (SPSS, 1998).

Some effort was therefore made to use an equal number of participant responses for each condition and within each age group. Thirdly, Gardner’s original study (1972a) had suggested that female participants might perform at higher levels than his male participants. Therefore should this effect be replicated in the present experiments, a substantial imbalance of male / female responses could distort the final results.

Secondly, in common with previous studies a score of 1 was given to each response where the participant had correctly identified the musical excerpts as being from the same or from a different piece of music. Each participant received three scores namely: total classical score, total popular score and total overall score. Participants

could obtain a maximum total overall score of 20 and a maximum score of 10 for the classical pairs and 10 for the popular pairs.

4.2.9 Method of Analysis.

The decision was made to use analysis of variance with post hoc analysis of means using either Scheffe or Tukey HSD tests throughout. There were a number of reasons for this decision. Firstly, the selected method of data analysis was required to not only highlight significant differences between the resulting means, but also to pin point any significant interactions between the two groups of classical and popular results. Secondly, all previous studies in this area had incorporated ANOVA as a means of analysing the resulting data. By following this tradition, it was therefore possible to more clearly compare and contrast the final data with those obtained in previous studies.

4.2.10 Practice examples

The test was always preceded by the two practice examples. Some thought was given to several issues relating to these practice examples, such as the actual musical examples to be used and also to any positive or negative side effects either the test examples or test procedure could have on the participants. Some discussion with one previous author of a similar study (Castell) took place and the type of music used and the possible effects of both the examples and the procedure were discussed. On balance and having taken advice from a previous writer the following procedure was adopted.

4.2.11 Practice examples: music selection

First, the two practice examples did not use any of the eight styles used in the test proper. This decision was made in order to minimise the chances of any possible confusion with the actual test pieces themselves and also so as not to give a possible 'more familiar' effect to any one particular style over the others. There was also a desire to avoid introducing too many new composers of one style, as this had been a possible limitation of the Tafuri et al. (1994) research. One pair was 'classical' in style while one pair was 'popular' in style. The two examples had a high level of contrast

and featured one 'same' example and one different example. The first 'same' example contained two excerpts from a modern, electronic sound with a Latin American beat by the Italian popular artist, 'Zuchero'. The extract contained modern synthesisers, electronic guitar and effects and yet was put alongside a slow Latin American rhythm usually associated with a very different style of music. The second different example consisted of an excerpt from a film soundtrack by Morricone featuring a panpipe dominant melody played above a symphony orchestra with a full range of instruments, although not all instruments were playing at once. The contrasting piece was taken from 'Music for Airports' by Brian Eno. This piece was slow, contemplative and featured no strong beat with all acoustic instruments.

These pieces were chosen because the style of each piece, at least in terms of what could be gleaned about a 'style' from the short excerpts, contained some elements of fusion with another style. Therefore it was felt that the participants might not react as strongly (positive or negative) to any of the pieces as they might to a more easily attributable style about which they may already have a strong positive or negative preference. It was felt to be quite hard to place the excerpts into a particular style and therefore it could well prove equally difficult to judge whether or not to react to it because of social influences or peer pressure. The examples contained one example of a 'same' and one example of a 'different' musical example.

4.2.12 Practice examples: marking procedure

Some thought and discussion was given as to whether or not to let the participants know how they had performed on the examples. Bridgeman (1974) reported that participants who had experienced 'failure' prior to the test situation appeared to perform less well than those participants who had experienced success. It was felt that some participants who may have got both examples wrong might have become negatively disposed to the remainder of the test. Diener and Dweck (1978) noted that children who began to experience failure in tests of non-musical skills showed very different behaviour towards the remaining test time than those who were experiencing some success. Those participants experiencing failure appeared to 'ruminate about the cause of their failure..... and spend little time searching for ways to overcome failure' (p.460). Participants were observed to engage in a whole range of negative behaviours

including devaluing the task; expressing negative attitudes towards the test; altering rules of the test and making negative predictions about their own personal performance. In contrast those participants who felt some success 'responded to "wrong" feedback chiefly as information leading to problem solution, not as a failure or as a prediction of future failure' (p.460). The most notable feature of the research showed that participants only exhibited these distinct behaviour patterns during an evaluation situation. Studies by Dweck and Leggett (1988) showed no correlation between their pre failure ability and the behaviour and ability they displayed following the failure condition.

With specific reference to music testing, O'Neill and Sloboda (1997) carried out an exploration into the extent in which post failure confidence influenced '.... recovery from failure during and evaluative achievement situation involving music' (p.20). Their findings suggested that the performance of some participants in a musical testing situation was affected after "one exposure to one brief failure situation" (p.32). Covington and Omelech (1979) suggested that the failure effect was greatest in those participants who either experienced failure following a high level of effort on their part, or in situations where praise was given for effort and yet the participant was aware of their own failure on the task. These authors point out the contradictory effects on students, which may arise through teachers' encouragement of achievement through effort. They note the subsequent behaviour strategies that participants may develop in order to avoid low achievement through maximum effort situations. As with the O'Neill and Sloboda study, Covington and Omelech suggest that the effects of even a small-perceived failure experience can be exacerbated when put in conjunction with a high level of effort.

As a consequence of these previous studies and as the pilot studies had suggested that a good level of sensitivity was taking place, it was decided not to give any right or wrong answers to the practice examples. By requesting a show of hands as an indicator of how participants had performed, it was possible to gauge whether or not participants appeared to understand the task. This was supported by the experimenter walking around the class to observe a good number of response sheets.

The relevance of the foregoing research to this study is that it contributes to the rationale for some of the principal features of the methodology, the musical examples and the assessment process. Each of the following experiments contains a more focused look at some further research that is pertinent to each specific and individual enquiry.

Chapter Five

Experiment One

Effects of lesson context

5.1. Background

This experiment investigated one aspect of change in the listening environment. Research (Martindale, 1990; Hargreaves and North, 1997, North and Hargreaves, 1996) has suggested that listener responses to music may vary considerably according to the immediate environment or the context in which the music is heard. Further research (Forsyth, 1975; Greer, Dorow, Wachaus and White, 1975; Murray, 1975) has argued that children may behave differently in music lessons as compared to their behaviour in other class lessons due to a number of intrinsic features of music as subject matter. This study investigates whether or not participants who experience the style sensitivity test as a musical activity within a music lesson, perform differently than participants who experience the test in a regular lesson. This will provide an interesting contrast to previous studies of style sensitivity which have always been carried out in whole or small class groups during a regular lesson or in small groups withdrawn from a regular lesson.

This experiment involved two conditions. In the first condition, participants experienced the style sensitivity test in a non-music lesson. The test took place in an area other than a music room or designated music space. Participants experienced the style sensitivity test either in whole class groups or in small groups withdrawn from a non-musical activity. This condition was subsequently labelled the control group. In the second condition participants experienced the same test within a music lesson or within a designated period in which they would expect to receive some form of practical musical activity. The physical environment was always a music room or music designated area. The design and procedures for the test were as noted earlier and the musical examples used were recorded and selected in the manner already described.

The hypothesis was that participants who experienced the test within the social context of a music lesson would show increased performance in style sensitivity, as compared to the control group in the non-musical environment. There were a number of possible reasons why this change in the test environment could produce significant differences in participant performance. Firstly, it was anticipated that the participants who experienced the test in the context of the music lesson should become more 'open eared' (Castell, 1982), that is more listener tolerant to all or some of the musical samples and therefore perform with greater accuracy. A period of listening to a wide variety of musical styles was a regular feature of music lessons in all the schools used in this experiment. This period of listening was always followed by a further musical activity based upon the excerpt the pupils had just heard. These musical activities often included open-ended questions about the mood of, or personal reactions to the piece or closed questions about the instrumentation or the tempo of the piece. Often further details regarding historical or biographical details of the composer or the piece itself were given and musical components such as form, texture or dynamics were discussed. Pupils therefore came to the music lesson with some expectation that a listening activity upon which subsequent activities would be based, may form part of the lesson. It was therefore anticipated that participants within the music lesson would listen to all the musical examples used in the test with more care and attention and not quickly dismiss those styles, which were beyond their immediate preference.

Secondly, one factor that may have affected the accuracy of responses in some of the previous studies was that some children resented having to miss a favourite or more popular lesson in order to listen to 'extra' music. Their already poor listener tolerance towards some of the musical styles used in the test may therefore have been further exacerbated by the imposition of 'extra music activity'. Thirdly, the phenomenon of *prototypicality* may exert an influence. This effect is discussed more fully in a subsequent section.

5.1.1 Listening environment

Castell (1983) had previously argued that the performance by older participants in her style sensitivity tests could have been affected by reduced listener tolerance. This decline in open earedness could have influenced participants' performance in a

number of ways and the implication made by Castell (1982, 1983), although not stipulated exactly, is that lack of tolerance of those styles beyond their immediate 'likes' were dismissed and therefore not fully attended to. This view received marginal support through participants' comments on the popular styles. Examples given included, "too slow" and "it was rubbish" (p.25). As argued earlier, this is the view that is further explored here; that the decline in 'open earedness' caused participants to be less tolerant of those styles beyond their immediate preferences and therefore the age effect was caused through a decrease in attentive listening. That is, participants dismissed those styles beyond their immediate 'likes' and did not attend to the excerpts.

As previously stated, it is probably very difficult to separate age effects from stylistic tolerance / musical style effect. It is probable that the age differences could be a result of stylistic tolerance changes with age, and / or their differential familiarity with the styles. However, the results of this experiment may provide some further evidence on this matter. A lack of statistical significance in the difference between the classical and the popular pairs might indicate some support for the idea that older participants show less tolerance for and therefore do not fully attend to those styles beyond their immediate 'likes'. This would outline an area of future research.

A number of studies have explored the effects of the immediate listening environment on listener tolerance, listener preference and listener perceptions. North, Hargreaves and Heath, (1998) carried out a study of participants' estimation of time spent in a university gymnasium as influenced by fast tempo and slow tempo music. Fast and slow popular music was played within a gymnasium on two consecutive days and participants were asked to estimate the duration of time they had spent in that environment. Their findings suggested that participants' perception of time and therefore their estimation of the duration they had spent on one activity could vary as a result of tempo. North and Hargreaves (1996) further investigated the effects of background music in a university dining area. They discovered that the participants willingness to remain within the environment and return to it were associated with the low, medium or high level complexity of music being played in the listening environment. Participants were more likely to rate the listening environment in a

positive way and make more use of an information stall when the background music played was of moderate complexity. Either low or high complexity background music was more likely to cause participants to rate the listening environment in a negative way and wish to leave it more quickly.

Konecni (1982) discovered that certain events occurring before the participant came into the test environment, did appear to exert an influence on the participants' preference for and tolerance of various musical styles. He discovered that participants having left an environment of high arousal (having experienced a high level of stress for example) or having recently been engaged in a complex activity or task, displayed poor levels of tolerance for complex musical excerpts.

Therefore it seems that listener reactions to and judgements about musical excerpts may well be the result of a whole set of interactions between the listener, the listening environment, the required task and style of music. Castell may well be correct in her implication that peer group sub cultures increasingly shape the *overall* publicly declared musical style preferences in young adolescents. These clearly defined and highly significant preferences for some musical styles must therefore, also define those styles to which participants have *overall* poor levels of tolerance. However, as reported above, those styles or excerpts of music to which listeners have overall high or low levels of tolerance may change within the confines of, and relatively brief context of another listening environment such as a music lesson. Here, listener tolerance and preference may be subject to influences that are far more localised, immediate and transient. It may well be that within the style sensitivity test, listeners exhibit low levels of tolerance for musical excerpts in the classical style, but this may be due to a lack of tolerance for a more complicated style of music than an overall high preference / low tolerance of musical styles due to social influence. For example, although some control was possible over the parameters of the test environment, no control was possible over the task or activity from which the participants came immediately prior to the test. Based on the findings of the research detailed earlier (Konecni, 1982; Hargreaves and North, 1997), participants who had been timetabled for maths or science, or any activity involving more complex and possibly stressful tasks and activities immediately prior to the test, may well have had lower levels of

tolerance to the more complex musical examples. Likewise, greater levels of tolerance for more complex musical excerpts may be expected from participants having experienced a quiet practical activity immediately prior to the test situation.

In summary, the overall declared musical preferences / tolerances of participants that may in fact be socially influenced may also be subject to more immediate, short lived and environmental influences.

5.1.2 Prototypicality

Authors such as Martindale, (1990) and Martindale, Moore and West, (1988) have explored the effects of music appropriateness to environment or *prototypicality* on listener preference. Martindale has argued that at least for naïve audience members, “meaning is by far the most important determinant of preference” (Martindale, 1990. p.42). He suggests that listener preference, and therefore possibly listener tolerance is greater in cases where naturalistic settings incorporate a musical stimulus which is more typical of the music usually played in that particular setting. That is, listener preference is greatest when the actual music heard matches their expectations of which type of music should be played in that setting. For the participants within the setting of the music lesson, it was 'typical' for them to be exposed to a wide range of musical styles. It was typical to hear both popular and classical musical examples. It was also typical to hear and further explore in a variety of ways, styles that were not necessarily a 'preference' of the listeners. Accordingly, within this experimental condition, participants could become more tolerant of the musical examples used in the test, as it was both appropriate and typical for that environment.

5.1.3 Test environment

A number of researchers have explored the effects of the test environment on participants' performance. Riecken (1962) suggested that very little was known about the physical environment in which experiments took place. Mintz (1957) discovered that participants' ratings of negative images of faces for pleasantness were influenced by how pleasing the test environment had been made prior to the test. More pleasant surroundings produced more positive ratings whilst unpleasant surroundings produced more negative responses. Rosenthal (1976) noted that the physical characteristics of

the test environment could affect both the behaviour of the experimenter and the participants. In his experiment, participants were required to rate photographs of faces according to the degree of success they felt the individual in the photograph had recently experienced. This “success estimating” test was operated by a total of 14 different experimenters. Each experimenter was randomly assigned to one of eight different laboratories in a university psychology department. The rooms were labelled as either ‘professional’ or ‘disordered’ environment. The ‘disordered’ environment was more natural in that participants experienced the room exactly as it would have been for lectures. The rooms were untidy and in a disordered state. The ‘professional’ rooms received some changes prior to the experiment. Each room was tidied, arranged in an orderly fashion with good lighting. Rosenthal described them as being ‘more comfortable’. Interviews with participants following the test experience suggested that participants had observed some differences in the behaviour of the experimenter according to changes in the room environment.

The main finding of the study was that participants who experienced both the test and the test presenters in the more ‘disordered’ laboratory were more serious about the purpose of the test than those who presented the test in the more comfortable and orderly environment. Rosenthal speculated that the reason for this result was that the disordered laboratory was more typical of the normal environment in which scientific studies are usually carried out. The room matched the stereotyped environment of both the experimenter and the participants. Therefore the test itself was seen as being far more genuine and to be taken far more seriously because it was being carried out in the most genuine and appropriate place. It may well be that in taking the test within a music lesson context, participants will perceive the musical style sensitivity test in a more genuine and more business like way. It may seem a more bona fide activity such that all excerpts will be taken more seriously.

5.1.4 Musical environment and testing

The majority of studies exploring the effects of the test environment on performance have been carried out using tests that contain non-musical material (such as I.Q. tests) and contain no specific reference to musical environments. However, a number of studies have explored these aspects purely in connection with music lessons. Kuhn

(1975) explored teacher approval / disapproval effects on 5th. grade participants in music lessons. The study explored the effect of three conditions (approval / disapproval; neutral interaction and non contact) on participants' social behaviour, musical achievement and attitude. The results suggested that participants receiving approval did increase the amount of time they were observed to spend 'on task'. No significant differences were observed between all four groups in terms of attitudes towards musical activities. No significant improvements were observed in musical achievements between any of the groups.

Forsyth (1975) replicated Kuhn's methodology to explore the effect of approval / disapproval on student attentiveness with music teachers and classroom teachers. The results suggested that the same participants appeared to be far more attentive in music lessons than in the other classroom lessons regardless of the level of teacher approval / disapproval. It was also suggested from the results that music teachers demonstrated a higher level of social approval responses than other teachers. The study suggested that children behave differently and react to teacher responses differently in music lessons than they do in other lessons. This led Forsyth to suggest that: " this finding suggests that a subject matter variable may relate to the results. Music appears to be 'naturally' reinforcing" (p. 49). Whilst the study suggests that children do behave differently in music lessons, there is no clear evidence as to the cause of this. Some components within the subject matter and music teacher behaviour patterns appear to interact and influence participant behaviour.

Greer, Dorow, Wachaus and White (1975) found that participants in music lessons showed preferences for the style of music taught under high approval conditions from a teacher. These authors suggested that music teacher 'approval / disapproval' behaviours are strongly associated with students' acquisition of music discriminations, such that, "teacher - generated approvals are potential reinforcers for students' subsequent music selections" (p.110). They suggest that teacher reinforcing behaviour has three possible outcomes in student behaviour. First, student avoids similar music in the future; second, student increases time spent with similar music in the future; or third, teacher does not influence student listening behaviour. In their study, participants received musical instruction from experts in the fields of jazz, rock

and classical music. Participants were also given the choice of listening to a variety of styles of music, (rock, jazz, electronics, classical music or white noise) under either high or low adult approval conditions. The findings suggested that participants' listening choice was affected by the amount of approval / disapproval ratings given by the 'expert' during the period of musical instruction. Participants taught a style of music under high approval conditions opted to listen to far more of that music. However, in the long term, personal listening preferences did not change.

5.1.5 Summary and background literature

In summary, the purpose of the experiment was to explore whether setting the task in the context of a music lesson would affect the tolerance of listeners to all or some styles of music. Castell (1982) had suggested that pupils' scores on both popular and classical pairs may have been affected by social influences. She argued that listener tolerance to some popular music samples could be affected by the clearly defined and narrow musical preferences that existed amongst the 13 and 14 year old pupils. These preferences caused participants to dismiss and therefore not fully attend to those styles beyond a small number of acceptable styles. Our review has shown that listener tolerance may be affected by both the typicality of the musical samples to the test environment (Martindale, 1990), and also the suitability of the test and the test materials to the testing environment itself (Rosenthal, 1976). Furthermore, it has been shown that pupils often behave differently in music lessons than in other regular lessons because of the natural reinforcing nature of listening to music (Killian, 1990; Kuhn, 1975).

Thus, our review suggests that listening environment, and in particular a music lesson environment may increase listener tolerance to either one or both of the musical styles used in the sensitivity test. In the classical pairs it is expected that participants will achieve higher total scores when receiving the test in the music lesson than those participants receiving the test in the regular lesson. This may be because participants in the music lesson are more prepared to attend more closely to the classical samples because;

- The classical samples are more typical of the music usually associated with the music lesson (Martindale, 1990);

- That the physical environment of the music room somehow reinforces the validity of the test (Rosenthal, 1976)
- Participants are not resentful of having been withdrawn from a more favourable lesson
- Participants may behave differently within the music lesson setting (Killian, 1990).

In the popular pairs, participants in the music lesson may perform equally well across all four popular styles given points 2, 3 and 4 given above. The main expectation of the results of the experiment is that amongst the participants in the music lesson, scores in the classical pairs will be equal to or more closely aligned with the scores achieved in the popular pairs, i.e. that the significant differences found in the previous four studies (Gardner, 1973a; Castell, 1982; Addessi et al., 1996; Hargreaves and North, 1997) between classical and popular pairs will not be present.

5.2 Design and procedure

The experiment incorporated two conditions, C and E. Condition C was a control group. Participants received the same style sensitivity test during a period when they were not expecting a music lesson or a musical activity; as was the case in other previous studies. The participants received the test in their classroom setting and the test either interrupted or replaced their usual lesson. Participants in condition E experienced the test within a period during the school day that had been timetabled for music. This means that the children came to the test expecting to take part in some form of musical activity from national curriculum designated schemes of work. All tests in both conditions were carried out during an afternoon period in the school. In the opinion of the teaching staff, physical education and games were regarded as particularly favourite lessons of a large number of participants, particularly male participants. Therefore no test replaced or interrupted a physical education or games lesson.

The material in the musical style sensitivity test and the method of introduction was used according to the procedures already described in Chapter 4. All participants in

both conditions experienced 20 musical examples (10 classical and 10 popular pairs). The order in which the examples appeared in the test was identical in both conditions. The order in which the musical excerpts appeared was selected according to the criteria already described in a previous chapter.

Following each test, the papers were checked through. Any papers where the participant had clearly not understood the task, for example, where both the 'same' and 'different' answers had been circled, were withdrawn. Following this initial check, either the class teacher or a member of the music department withdrew an appropriate number of responses, as described earlier, in order to reduce the total number of responses to 40 in each case.

5.2.1 Schools

The experiment involved responses from 320 participants comprising of 80 children in each of four age groups namely 7/8 years; 10/11 years; 13/14 years and 15/16 years. In both the 7/8 and 10/11 year age groups, participants were taken from two different schools: a small village primary school in West Sussex (32 participants) and a much larger Middle school in a South London Borough (61 participants). This was done in order to achieve a wider social mix. Gardner's original study had incorporated only a small number of participants from a predominantly white, middle class university town in Massachusetts. Gardner commented that this narrow social mix and limited research population may limit the generalisability of the results. Therefore in the second study, Castell had deliberately incorporated participants from a wide range of ethnic and social backgrounds (as denoted by main household occupation). The individual schools differed in catchment area; the number and variety of musical activities that were offered by the school and the amount of music and musical activities that the children experienced within the average school week. Both of the schools described themselves as offering some form of musical appreciation and one school operated a choir, a non-selective orchestra and a variety of recorder groups. The second school offered small recorder groups and individual recorder lessons to those children who were interested. Some of the participants received some group tuition on either violin or flute through a visiting teacher. Both schools offered some musical activities over and above timetabled music lessons.

In the 13/14-year and 15/16 year old groups, participants were taken from a large comprehensive school in South London. Class groups were mixed ability (established by SAT results) and contained participants from a number of different cultural and social backgrounds. The small music department taught National Curriculum aims and objectives for music and pupils had the opportunity to take part in a number of extra curricular music activities including choir, swing band and an annual musical. A small number of the participants (approximately 15%) took advantage of these activities.

5.2.2 Participants

7/8-year age group.

A total of 93 children from two schools were used in this age group. 47 participants were tested in the non-music lesson - condition **C** (control) and 46 participants were tested in the music lesson – condition **E** (experiment). In each case, the class teacher randomly removed the appropriate number of response sheets to make equal numbers of 40 responses for each condition. No class group involved had the music teacher for their own class teacher. One third of the participants in each condition experienced the test in a small group whilst two thirds experienced the test in a full class group. The music lesson condition was operated within the music room.

10/11-year age group.

There were a total of 92 participants in this age group and 46 participants were tested in each condition. These results were reduced to 80 by random withdrawal.

13/14 year age group and 15/16 year age group.

All participants in both conditions experienced the test in full class groups. In condition **C**, the participants experienced the test in a period normally timetabled for Design or Food Technology. There were a total of 94 participants in the 13/ 14-year-old age group. Condition **C** had 46 participants and Condition **E** had 48.

In the 15/16-year-old age group, there were a total of 88 participants in both conditions. Condition **C** had 43 participants and Condition **E** had 45 participants. All participants were taken from one full year group of six classes. The designation as to which groups received the music lesson condition and which groups acted as the non-

music control group was decided purely on grounds of timetable within the school. This was deemed to be random as the tester had no influence over which groups were present at the time of the test. The school assessed all groups as being parallel mixed ability classes. Participants' answer sheets were then score. A maximum total score of 20 was possible with a maximum score of 10 possible in each of the classical and popular pairs.

Previous studies (Castell,1982) had found no effect on participants' performance in style sensitivity tests between those participants experiencing the test in small groups and those experiencing the test in full class groups. In the present study, participants in the music lesson variable experienced the test in an identical way to that in which they experienced their usual music lessons i.e. either in small or full class groups.

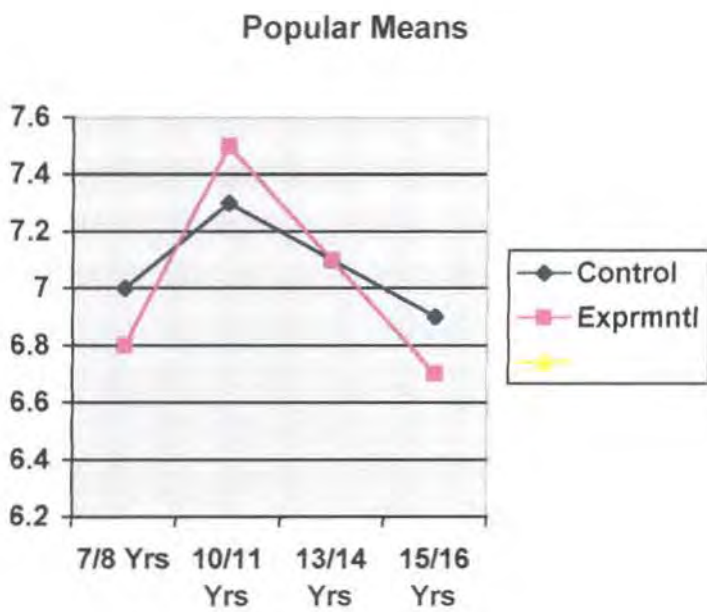
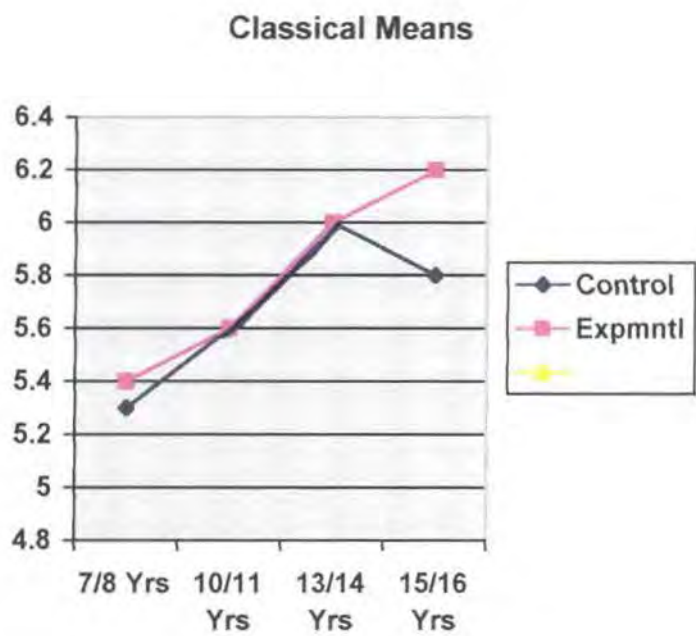
5.3 Results

A three way mixed analyses of variance was carried out on the style sensitivity scores with one within-participants factor (musical style: classical / popular extracts) and two between participants factors - age group [4 levels] x condition. As all participants experienced both the classical and the popular pairs, this was a repeated measures analysis. Further analysis of both the individual classical and popular score means was carried out by the use of post hoc Tukey HSD tests.

The main effect for age was statistically significant ($F [3,312]=3.68, p = .01$); and likewise the main effect for style was also found to be significant ($F [1,312] = 210.37, p=0.00$); whereas that of experimental condition, was not significant. ($F [1,312] = 0.01, p= 0.91$). For the three main two - way interactions, there was found to be no interaction between condition and age ($F [3,312] = 0.12, p = 0.95$); nor between style and condition ($F [3,312] = 0.69, p = 0.41$). However the interaction between age and style was significant ($F [3,312] = 5.87, p=0.001$). Lastly, the interaction between age, condition and style was not found to be significant ($F [3,312] = 1.11, p=0.35$)

The interaction means for the classical and popular scores are plotted below in **Fig. 5.1**

Figure 5.1



5.4 Discussion.

The first observation from the results of this experiment suggests that in common with the previous studies, the overall standard of accuracy in both popular and classical pairs was high with even the youngest age group achieving scores above those expected through chance. This appears to support Gardner's initial statement that even young children are capable of discriminating cues in musical stimuli and making reasoned judgements about their probable source. Secondly in common with Gardner and Castell's results, age proved to have a significant effect. 10/11-year-old participants achieved the highest scores in the popular pairs in both control and experimental conditions; that is scores higher than both the older groups. In the classical pairs, the 13/14-year-old participants achieved the highest scores in the control condition but the 15/16-year-old participants achieved the highest scores in the experimental condition.

However, the results were not in keeping with those obtained by Addessi et al. (1995), nor those obtained by Hargreaves and North (1997) who both found that the main effect for age did not reach statistical significance. This anomaly is surprising in view of the fact that the same classical pieces were used in all three experiments and the same popular pieces were used as in the Hargreaves and North experiment. Hargreaves and North's results did however show that the mean scores for the classical and popular pairs did increase with age, but did not reach statistical of significance. So the lack of agreement between these two findings may simply be one of degree. In the present study, only in one category did the 15/16 year old participants perform better than 13/14 year old participants; that of classical pairs in the musical environment. Again this trend of less accurate responses in the older participants was in keeping with the previous studies of Gardner (1972) and Castell (1983).

The main effect for style, in common with previous studies, was found to be significant. That is participants performed more accurately on the popular pairs (familiar genres) than the classical pairs (unfamiliar genres). The expectation was that participants in the music lesson would respond to the more typical and authentic musical environment and therefore increase their level of performance in the classical

pairs bringing their classical and popular scores closer together. This supports the finding that the main effect for experimental condition was not found to be significant. Participants' did not perform more accurately in the music lesson.

A further analysis of the classical and popular means separately using Tukey HSD tests did show a number of differences between the means of the four age groups. In the popular pairs, the mean scores of the 10/11-year-old participants were higher than that of the 7/8-year-old participants ($p=0.05$). This suggested that most improvement in performance occurred between the ages of 7/8 years and 10/11 years. The mean scores between the 10/11-year-olds and the 15/16-year-old participants showed a decrease in performance that also reached a level of statistical significance ($p=0.02$).

In the classical pairs, the difference in performance between the 7/8-year-old participants and the 13/14 and 15/16-year-old participants was significant, ($p=0.01$) in both cases. The responses suggested that participants' sensitivity to classical musical styles may follow a developmental trend that is very different from that of sensitivity to popular musical styles.

Within the popular pairs, the 10/11-year-old participants appear to be knowledgeable, attentive and able to identify accurately pertinent features in the musical styles whereas the older participants either become more discriminating and sensitive to more discreet differences within the musical stimulus; or more selective in those styles of music they are willing to tolerate.

The main hypothesis of the experiment was that a change in test environment would affect the performance of participants in a style sensitivity test. This hypothesis was not proven and there are a number of reasons why this may not have happened.

Firstly, some artefactual element within the test design or the musical stimuli used may have prevented this effect from being identified. Secondly, of all the possible social variables that may exert some influence on the participants' response to the musical stimuli used in this study, the act of including the test within a musical environment was not seen as significant by the participants themselves. The status of music, music education and music lessons within the schools themselves may, in

some way be a determinant of this and this deserves further investigation. Third, the instrument of measurement lacked the required sensitivity to measure this particular variable.

5.5 Conclusions

As with previous studies, the main effect for age was found to be significant; style sensitivity did increase with age of participant. The fact that the main hypothesis was not proven however deserves some comment. There are a number of possible reasons for this.

- Some artefactual element within the musical samples used exerted a strong influence on participants' responses.
- Interactions between the test and participants' attitudes towards music lessons within the individual schools. For example, pupils in individual schools may have viewed music lessons in a negative way; or music was viewed in a very positive way and pupils resented the test replacing their more regular and typical musical activities.
- Some limitation within the test design or some effect that remained undetected. For example, it may be that the music lesson environment did affect participant performance and attitude but this resulted in a behaviour change beyond the sensitivity of the test design.
- It may be possible that the music lesson environment did affect the participant performance and attitude, however some further influence operated within the control group resulting in artificially high results from those participants. Similarly, it is accepted that school itself is a very powerful and influential environment in determining behaviour. Findings may have been found to be different if a further condition had involved testing in a social context beyond the school environment.

Taken as a whole, the results appear to add support to the idea that participants between the ages of 8 and 14 years appear to show far higher levels of discrimination when judging musical styles with which they are familiar. The results also appear to

support the idea that a number of social and cultural factors exert considerable influence on the determinants of style sensitivity and teasing out individual influences may not be possible. The subsequent experiment attempts to explore one of these influences, that is the possible influence of the test presenter.

Chapter Six

Experiment Two

The effects of variation of the test presenter

6.1 Background

In experiment one, the same presenter operated a musical style sensitivity test in two different conditions, one test environment was musical whilst the other was not. In the conclusion to that experiment, it was suggested that the test presenter might have exerted some influence on the participants' performance. This experiment attempts to explore this issue further by investigating whether or not the age, cultural style and gender of the presenter can influence style sensitivity.

This second experiment incorporated three experimental conditions. In the first condition, participants received the writer as the presenter of the test, as was the case in experiment one. This group were designated as a control group and are subsequently labelled as Condition **Co** (control). In the second condition, the participants experienced the test given by a young female test presenter, subsequently labelled Condition **Ef** (Experimental: female) and in the third condition the participants experienced the test from a young male presenter, subsequently labelled Condition **Em** (Experimental: male). In total, twelve tests were run, one test in each of the three conditions in each of the four age groups.

The experiment sought to explore further the social context in which the musical stimuli were heard. It was hypothesised that participants in conditions **Em** and **Ef** would perform significantly better than participants in condition **Co**. It was thought that participants might become more listener tolerant to all or some musical styles when presenters who were closer both in age and culture presented the test.

Chapter 3 detailed a number of studies that have attempted to investigate and define those attributes of a test presenter that may produce observable effects on participant performance. In addition to those detailed earlier, the possible influences of the race,

gender and age/experience of the presenter have all been systematically investigated and a summary of the findings is presented here.

6.1.1 Race

A substantial amount of research has been carried out on the interaction between the race of participants and the race of the experimenter, as evidenced in participant accuracy in a variety of tasks under test conditions. Researchers found varying results in studies involving black and white participants when presented with tests by black and white presenters. Although many of the studies reviewed by Sattler (1970) proved inconclusive, in summary it was suggested that white presenters could affect the performance of black participants in measures of task performance. In tests that involved intelligence testing, the data remained inconclusive (p.156). Black presenters did not appear to affect the performance of white children in quantitative test scores, they did appear to affect their responses in more open-ended tests (Gahagan et al.1969).

Results from studies using the WISC test material also proved to be conflicting or inconclusive. Moore and Retish (1974) found that black participants performed better with a black presenter; Samuel, Soto, Parks, Ngissah, and Jones (1976) found that both black and white participants did better with a white examiner and Samuel (1977) found no significant difference. Kennedy and Vega (1965) discovered a number of interactions between black / white experimenters and black / white participants. Although some increase in performance in an oddity discrimination task was observed in black participants with a black examiner, the main significant differences emerged during further testing involving verbal praise, verbal blame and verbal control conditions. Verbal criticism and praise from a black experimenter increased black participants' performance whilst verbal criticism from white experimenters caused black participants' performance to decrease.

Further studies by Phillips (1966) investigated the effects of the gender and race of the test presenter. This additional condition of gender produced further interesting interactions. It appeared that participants' domestic environment influenced the extent to which participants were influenced by these two particular presenter attributes. Phillips discovered that race and gender of experimenter could vary as a function of

whether or not the participant had a father living at home. Young black males with no father living at home appeared to be more affected by both the race and gender of the presenter and performed better in all conditions involving black male presenters.

Many of the reported research papers have been criticised as having weaknesses in the methodology. (Sattler, 1970; Lutey and Copeland, 1982) Amongst the main criticisms were:

- a) Unequal numbers of black and white experimenters were incorporated in the studies.
- b) Lack of random assignments of participants and experimenters to conditions.
- c) Inappropriate statistical analysis of the resultant data;
- d) Lack of control of other variables such as experimenter expectancies, age, gender, stature of experimenter and culture; that is black experimenters have included Asian, Negro and Indians.

Summers and Hammonds (1965) demonstrated that race of experimenter did affect how participants felt they should respond. More participants claimed to be racially prejudiced in an anonymous questionnaire when given by two white experimenters than when it was given by one white and one black experimenter. Likewise Katz, Robinson, Epps and Waly (1964) observed that in a test involving rhythmic co ordination, black participants performed better with white experimenters by being more willing to demonstrate their good sense of rhythm, which they believed to be an expectation of the experimenter. When the same test was restructured as an intelligence test, black participants performed better with a black experimenter. Within this present experiment, race was not an issue. The majority of the participants were white (approximately 98%) and all presenters were white. However, Lutey and Copeland (1982) attempted to explore the issue of race further. In circumstances where race appeared to be exerting some influence on participant performance, they attempted to identify more precisely which experimenter attributes were causing the interactions. They argued that 'race effect' could operate in three individual and separate ways. Firstly improved scores amongst black participants with a black experimenter could be explained in terms of 'kinship'; that is an improvement in

performance through an increase in motivation. Secondly, the performance of black participants with a white experimenter may decrease because of a rejection of what the participants see as 'white values'; that is performance is affected by lack of effort / motivation. Thirdly the decreased performance of black participants with a white experimenter could also be more the result of the experimenters':

" failure to understand the subject's cultural milieu, including attitudes, values, cognitive style, language usage, behaviours in the testing situation, and other unique qualities" (p.131)

Labov (1970) suggested that many experiments were designed and set within a cultural vacuum. However, Labov was in fact arguing that cultural and race boundaries were one and the same thing. That is within his work amongst black participants, the linguistic ability tests given to black participants were based in a predominantly white culture by people of the white race. The culture and the race divisions were identical. However, in many populations, different cultures may exist within the same race and different races may exist within the same culture. That is, different teenage 'cultural sets' based around musical styles can exist within teenagers of the same race. Similarly, different races of teenagers can exist within one musical culture. That is within a musical setting, the boundaries between race and culture are not one and the same but can cut across.

Therefore the influence that has been termed 'race effect' by some researchers, at least in part may be as much to do with 'cultural differences' as they are to 'race differences'. Within Lutey and Copeland's analysis there are three categories of influence. Firstly, increased motivation to perform better through a feeling of kinship. Secondly, decreased motivation as a result of a rejection of the values of another culture within which the test is seen to be set and of which the experimenter is seen to be representative. Thirdly, a failure to understand the task through ignorance of the culturally bound rules implicit within the test materials or structure.

In terms of the first category, 'kinship' may be felt just as much between two teenagers belonging to the same street culture, as it can between two participants of the same race. Therefore the increased motivation to perform better on a test may

occur between participants of the same culture as much as it may occur between participants of the same race. Secondly, and similarly, the lack of motivation to perform adequately through a rejection of the values, ideals and expectations of a specific culture as bound up in and represented by the test materials and the test presenter can also occur equally between two participants of a different culture / same race; as it can between two participants of different race. Thirdly, Labov (1970) states that participants who are not part of the immediate culture within which a test is set frequently "fail to perceive and discriminate in ways which the tests are designed to measure" (p.12). An example of this would be that standard classically trained musicians could categorise music by discriminating between harmonies, forms, melodies, orchestration and tempo patterns, whilst a teenager may categorise music by discriminating between production techniques, electronic effects and the final instrumental mix. The style sensitivity used in this experiment is set within two particular musical cultures, one of them, classical, may well be a culture beyond the immediate street music culture of the participants involved within the test situations. The way in which musical discriminations occur and the way in which music is classified and therefore thought to be the 'same' or 'different', can be as different between the sub cultures of the same race as it can be between the cultures of different continents and races.

A further point on this issue is that a test may be seen as a cumulative experience. That is, participants do not necessarily react to each section/ question / task in isolation from the other parts of the test. In any response to any part of a test, a participant is possibly reacting as much to an early part of the test material as they are to the task at that moment. For example, within the present test material, participants are presented with a set of verbal instructions. They are then presented with a pair of musical examples. In order to establish if these two musical pairs are the same or different, participants are required to discriminate and categorise what they are hearing. If, for example the two pairs are examples of Ska and Grunge, then discrimination may take place along the lines of instrumentation, 'clean' guitar, horns and light bass contrasted against 'compressed' guitar, heavy bass and over modulation. This experience early on in the test may shape all subsequent discriminations. That is, the criteria used in discriminating and categorising an early

example may be applied to subsequent tasks within the test. Therefore, any participant who is strongly committed to a teenage sub culture that involves discriminating and categorising music according to a particular set of rules will not necessarily be able to 'perceive and discriminate in the ways the test' was designed to measure.

Within the context of this study, although 'race' was not seen as an issue, there was some difference between the teenage cultures of the participants and the presenter who may in turn have been seen as an extension of the school teaching culture. This experiment employed three different presenters, partly in an attempt to examine if some form of 'culture effect' could influence perceived stylistic sensitivity. To refer again to the three categories of Lutey and Copeland, it was questioned if an increased performance could be engendered through increased motivation brought on through the 'kinship effect'. Would participants in conditions two and three feel more 'kinship', and less inclination to reject some of the musical material by experiencing the test with the two younger presenters who may be seen as being culturally closer to participants, than with the older one. Alternatively, some participants may reject the values of a pronounced musical culture different from their own and strongly reject the two younger presenters.

6.1.2 Experimenter gender

Lutey and Copeland (1982) reviewed the results from 17 studies that had investigated the effects of presenter gender on participant performance. These studies showed a mixed set of results but provided some suggestion that experimenter gender could react strongly with other factors within the test. The results confirmed that male and female experimenters often achieve significantly different results from their participants although it is not always possible to predict for any given type of experiment just how participants' responses will be affected by the sex of the examiner. A number of studies involving intelligence tests (Morales, 1977; Hanley, 1978) found that there were no apparent effects of experimenter gender on participant performance, although authors such as Samuel (1977) found that female examiners elicited better results from both male and female participants. Yet further studies by other authors (Bradbury et al., 1975) found that male participants scored higher with female experimenters. Lutey and Copeland pointed out that:

"age and not just sex of the children influenced whether they did better with a male or female tester" (p.132)

Quereshi (1968) observed significant interactions between male / female participants and male / female experimenters as a function of participant age, whilst Bittner and Shinedling (1968) demonstrated that in tests involving the Piagetian concept of conservation, male experimenters could elicit better test performances from male third grade participants than could female experimenters. Pedersen, Shinedling and Johnson (1968) then carried out a further study on 24 third grade participants using the mathematics sub-test of the Wechsler Intelligence Scale. Their overall findings suggested that participants performed better in experimenter- participant same sex conditions. Male participants did perform better with male experimenters but female participants appeared more susceptible to this effect. Under experimenter / participant same sex conditions, the female participants improvement in performance was much greater. A further claim made by the study was that mathematical test content was less susceptible to extraneous variables e.g. gender of experimenter, than experimental materials involving language.

As a result of the findings of this study, the authors suggested that:

" even greater experimenter effects may occur with measurement instruments that are more sensitive to the effects of extraneous variables" (p.253)

The suggestion is then that the magnitude of the effect of experimenter gender on participants' test performance may vary according to the style and measurement sensitivity of the test material. Additionally, it may also vary between individual academic subject areas. The authors suggested that the improvement in test performance in male examiner/male participant/ mathematics conditions and female participant/ female experimenter/ English language conditions could be explained in terms of the greater number of male mathematics teachers and the greater number of female English teachers within their education system. They argued that this gender imbalance amongst participant teachers might have produced an 'educational artefact' specific to the community in which the testing took place.

Other studies, such as that undertaken by Cieutat and Flick (1967) found that participants performed better on the Stanford-Binet test under conditions where the presenter was of the opposite sex whilst Stevenson (1961) found the differences in results between same sex experimenters greater than the difference between different sex experimenters and that male examiners tended to elicit a greater overall performance from their participants. This however changed when verbal reinforcement of the participants' performance was introduced into the test situation with female examiners then obtaining greater performance from the youngest participants (3-4 years). In the older age groups (9-10 year olds), the trend was reversed with male examiners obtaining the greater performance. In the middle age group (6-7 years) participants performed better with an examiner of the opposite sex. Stevenson and Allen (1964) explored experimenter gender effect amongst older participants. 128 male/ female college students were tested using both female and male experimenters. Their results suggested that the greatest increase in performance was achieved through participants experiencing the test from an experimenter of the opposite sex.

It has therefore been shown that participants' reaction to, and their susceptibility to experimenter gender effects can vary as a function of academic subject, age of participant, test material and the behaviour of examiner. The question is whether or not experimenters of different sex achieve different results from their participants because they behave differently or because of their appearance. Friedman (1964) and Katz (1964) carried out studies of experimenter behaviour in relation to the amount of eye contact exchanged between participants and examiners. Their studies suggested that the interactive behaviour between both sexes of examiner and participant frequently changed throughout the various stages of the test. The main findings were that female participants received almost two and a half times as many glances from experimenters of both sexes as did male participants. During the reading of formal instructions, examiners exchanged twice as many interactive glances with participants of their own sex and significant differences also occurred in the number of smiles exchanged with or given to male and female participants. Interactive exchanges involving smiling were greater with female participants than male participants.

Sarason and Harmatz (1965) carried out a series of experiments involving the teaching of first person pronouns with male and female presenters exploring a number of behavioural variables. They found that participant performance appeared to be a complex mix of experimenter and participant gender, hostility of experimenter behaviour, and the degree of contact between experimenter and participant and the prestige of the experimenter. The results caused Rosenthal (1976) to suggest that:

“ It appears that at least in studies of verbal conditioning, when an experiment is so designed as to permit the assessment of complex interactions, these interactions are forthcoming in abundance. Only rarely, however, are most of them predictable or even interpretable” (p.43)

One further significant study was that done by Stevenson and Odom (1963). This study involved a lever pulling exercise, which included a reward. Male presenters achieved better performance from 6-7 year olds and 9-10 year olds than did female presenters. The significant feature of this research was that the presenter was not present during the exercise itself but left after the instructions had been given. Stevenson, Keen and Knight (1963) achieved similar results on a marble dropping exercise with younger children aged 3-5 years old.

This brief review illustrates the fact that the influence of and participants' susceptibility to experimenter gender is likely to be a complex issue affected by interactions between the age of the participant, the gender and behaviour of the examiner, the gender of the participant, the experiment material used and the nature of the test activity the individual was required to do. Rosenthal (1976) states that both simple and complex reactions to experimenter gender have multiple sources, “ including those that are genetic, morphological, endocrinological, sociological and psychological” (p.56). He concludes that within any formal testing situation, because of the different behaviour patterns of male and female examiners both passive and active, male and female participants may not psychologically be in the same experiment (p.56).

The findings of the studies reviewed in this section are of interest and are noteworthy but two considerations need to be added. Firstly, a large percentage of the reported literature has been carried out in a one to one situation. The relationship between the

experimenter and participant was therefore more intense and close than in a group-testing situation. Secondly, no further research has been done as to how experimenter gender effects operate in a variety of test situations, for example those that depend on personal interaction for their data collection and those that do not. That is, experimenter / participant interaction could be placed on a continuum with total interaction at one pole, i.e. in a one to one interview situation where the relationship between experimenter and participant is vital for the collection of the data; and a mailed questionnaire or self administered test at the other pole. No data exists as to whether or not experimenter gender effect varies under each of these extreme conditions.

The experiment reported here used three presenters, one female and two male. The research reported above does suggest that the gender of the presenter may create some interaction or affect the responses of some participants with some test materials. As the previous research was inconclusive as to the precise way and the exact circumstances in which gender may influence participants' test performance, careful consideration was given to the question of incorporating a 'control' measure for gender into the test situation. One such measure for example, could be to define and isolate further qualities in presenter attributes. This would create a large number of further varied conditions incorporating various permutations of presenter gender, presenter age and presenter culture. In practical terms, using two age groups of male and female presenter in just two cultural styles of dress on all four participant age groups could require up to thirty-two different research populations, eight presenters and over a thousand participants. The question to be addressed was whether or not the use of this substantial amount of human resources was appropriate.

There was no evidence to suggest conclusively that gender was the precise attribute of the presenter which influenced the test performance. As was noted earlier the interaction may be complex and affected by participants' age and gender; examiners' age, gender and behaviour; the type of test material used and the nature of the activity the individual was required to do. Furthermore there is no evidence to suggest that this style of test is affected by gender or by the cultural style of the presenter.

Similarly, there is no evidence to suggest that this test methodology is sufficiently sensitive to measure such influences, even if they exist.

It was therefore decided to run only three conditions and to see if the results suggested that further investigations on gender interactions would be worthwhile pursuing with this methodology. That having been said, a number of steps were taken in order to create some homogeneity of behaviour between the three presenters. Firstly, the two alternative presenters were made aware of those behavioural aspects in which male and female experimenters were often observed to be different. Each one read the literature review and noted ways (e.g. eye contact) in which their behaviour may differ. Secondly, all three experimenters witnessed the other two administering the pilot test in an attempt to produce some uniformity in test presentation.

6.1.3 Experimenter age

There have been a limited number of studies done which have tried to detail the effects of the age of the experimenter on participants' performance. Ehrlich and Riesman (1961) noted different responses in female participants according to age of experimenter. The overall responses suggested that age of experimenter could make some difference to the participants' performance but again, as with gender, the effect seemed to be a complex one involving other behavioural factors as well as age. Therefore no conclusion was reached as to whether or not it was age per se that created the effect. Therefore, similar comments apply as to those surrounding the issue of experimenter gender namely; there is no precise evidence that the style of musical test material used in the present experiment is influenced by experimenter age. Similarly, there is no evidence that this musical material is sufficiently sensitive to measure any interaction which does take place. To isolate 'age effect' further as a variable would require the same large amount of resources as indicated earlier in the section on gender. At this time it was therefore decided to remain with three initial experimental conditions.

6.1.4 Prestige effects

A number of studies have explored the 'prestige' effect. In these studies, researchers attempt to investigate the way in which the perceptions and judgements that

participants make about works of art can change as a function of the information they have about the work. The findings of this research suggest that the tolerance, preference and processing of musical samples may be affected by adding 'prestige' to a work. This can be achieved in a number of ways such as aligning the piece to a prestigious composer, or by raising the status of the performer or through majority consensus. A brief review of a sample of the literature is presented here.

Asch (1948) carried out some of the first reported studies and this formed the basis of a number of subsequent investigations in this area of psychology. In his first study, Asch re-evaluated and challenged some of the initial investigations on 'prestige effect', carried out by Lorge (1936). Lorge had provided participants with a number of written statements and attributed them to various authors; some more prestigious than others. Participants displayed a tendency to agree with or value the statements made by the more prestigious authors regardless of content. From this Lorge concluded that prestige acted arbitrarily on statements, regardless of its content or merit. Asch challenged this view and part replicated the study but included more detailed responses to the statements and their authors. Asch's study led him to propose an alternative thesis where "changes of evaluation require the transformation of content in response to altered contexts" (p.275). He argued that adding 'prestige' to a statement did not simply make the evaluation of that statement into an emotional one, but that it actually changed the whole cognitive character of the evaluation process.

The second study by Asch (1956) explored how participants appeared to be willing to suspend their own judgement in order to comply with a group decision. In this study, participants were asked to make judgements regarding the various lengths of lines. Asch found that participants were prepared to voice judgements that were clearly wrong in order to comply with the majority view of a small group. Radocy (1976) incorporating pitch and loudness matching tasks carried out a similar experiment, with similar results.

Much of the previous literature by Asch and others used non-musical stimuli whereas Duerksen, Radocy and others reported similar studies that more clearly explored the link between prestige effect and musical stimuli. Duerksen (1972) played student

participants two musical performances of the same piece. One piece was presented as being by a professional performer, whilst the second piece was labelled as being by a student performer. Participants consistently rated the professionally labelled performance as better than the student labelled performance. Radocy (1976) gave false information about musical samples to college students. Pairs of the same musical examples were played to the participants. There were two tests in three conditions. In test one, participants were asked to rate the better performance of the two examples and in test two, state their preference. In the first condition, no information was given regarding the individual musical performances. The second condition introduced a moderate bias. Prior to hearing the pieces, participants were told that one performance was by a professional whilst the second was by an amateur student. In the preference test, participants were told the first example was by a famous composer whilst the second was by a fictitious composer. In the third condition, labelled as "strong bias", participants were given strong indications about the choices supposedly made by fictitious listeners. The study caused Radocy to suggest that, to varying degrees, an authority figure did have some ability to bias preference judgements, "Who says or does something may be more important to a student than what is said or done" (p.119).

Chapman and Williams (1976) reported substantial differences in the liking for the same piece of music between three identical groups of participants. The piece, by Japanese modern composer Takemitsu, was played to three groups of adolescent participants. Each group was informed the piece was by a different composer. In the 'high status group' the piece was introduced as being from an LP by a member of the leading progressive group 'Pink Floyd'. Participants were told that Pink Floyd was one of, if not the best progressive group in the country at the moment (p.67). In the second 'low status group' group, the same piece of music was introduced as being by a 'leading Japanese composer of modern serious music'. Takemitsu was described as being held in high regard by critics of serious music. The control group were given no information about the piece of music they were about to hear. The results showed that participants tended "to rate the music more favourably when it was given high status" and that "the perceptual organisation of music varies with the context in which it is heard" (p.70).

Both the high and low status groups were presented with introductions to the piece that gave it status. Pink Floyd was described as possibly " the best" and Takemitsu as "highly regarded". The difference in responses within the two groups illustrated the point that any status or prestige attached to music had to come from or be associated with a source having credibility to the participants own teen culture. The third 'control' group was given no information about the music. Many of the participants felt unable to state exactly what their reaction to the music. A typical written comment was " I don't know what to think" (p.71). This type of response suggested that unless the music was placed within a context, participants found it impossible to make a value judgement on the piece.

Farnsworth and Beaumont (1929) gave participants false information about various aesthetic stimuli previously having been ranked highly by acknowledged experts. When asked to rate each stimulus for attractiveness, participants given the false information also gave higher ratings. Although not backed up by empirical evidence, Swanwick (1968) adds some anecdotal support to this. A small group of children were told that a piece of music came from the film, 'Fantasia' by Walt Disney. Following the hearing, the children were again told that it was a piece of music from the Sixth Symphony by Beethoven. The response from one child was:

"When you said it was Walt Disney, I wanted to like the music, because I like Walt Disney, and I did like it; but if you had said "Beethoven", I would not have wanted to listen to it" (p. 119)

This experiment seeks to explore the point made by Radocy, and supported by others, that it is who says or does something that may be more important than what is said or done. The musical test materials might be given some 'prestige' by the two younger presenters. The musical examples used are positively 'endorsed' by the presenters who establish a direct link to the music by stating the music is especially chosen as their favourite pieces. The experiment questions whether or not bias created through this form of 'prestige' can affect the level of participant's tolerance and consideration. As suggested earlier, this view is supported by Rodriguez and Webster (1997) who found their research supported the findings of LeBlanc (1984) who suggested that in terms of musical preferences, children demonstrated a gradual devaluing of the

opinions of their family and of their music educators and adolescents strongly favoured the musical preferences of their peer group.

6.1.5 Summary of background literature

Previous research has suggested that participants' test performance may be influenced, either singularly or by a combination of varying presenter attributes such as age, gender, race/'cultural kinship' and presenter mannerisms. In this experiment, three different experimenters present the same style sensitivity test to different groups of participants. The three presenters differ in age, appearance, gender, experience, dress style and possibly kinship. Previous research has demonstrated that each of these factors may, in certain circumstances and with differing test materials, affect the performance of participants.

Therefore to summarise, the hypothesis is that participants in conditions **Ef** and **Em** will perform with greater accuracy on the style sensitivity test than participants in condition **Co**. This may occur through one, or a combination of the influences reported earlier. In the present experiment, the large number of differences between presenters (age, gender, culture etc.) suggests that a large number of interactions might take place that may subsequently affect participant performance. Additionally, previous research into presenter attributes has employed test materials such as I.Q. tests, Stanford-Binet tests and Weschler tests. The present experiment employs musical test material and the purpose of the experiment is to attempt to establish whether or not any significant presenter / participant interactions can be observed through this form of test material. Determining exactly which, and the extent to which any of those presenter attributes influences participant performances is beyond the scope of this particular experiment. The main research questions for the experiment are firstly whether or not this particular form of test material i.e. the style sensitivity test, is a suitable and effective measure of varying presenter interaction effects and secondly it seeks to identify which of those varying presenter attributes appear to most affect participant performance .

6.2 Method

The experiment incorporated the three conditions already described. As detailed earlier, every attempt was made to ensure that all other variables, such as presentation and instructions were kept identical. All three tests were carried out during an afternoon period and followed identical timetabling considerations to those outlined in the previous experiment. Participants in condition **Co** experienced the test as presented by the writer. Participants in condition **Ef** experienced a second presenter who was a young female music student aged 19 years old. Participants in condition **Em** experienced the test from a third presenter who was a young male music student also aged 19 years old. Both students were recruited from a local university college and both expected to graduate as music teachers. Both students volunteered for the project and were selected for their pleasant manner, expressed interest in the project and other issues surrounding the psychology of music and lastly, for their prominent allegiance to current popular teenage culture as represented by dress code, style of hair and language. Both bore some resemblance to popular members of the current teenage popular cultures of popular music or of television soap opera. For the test situation the female experimenter presented herself with short dark hair with added coloured streaks of blue and a small amount of added metal work to ears and nose. Clothing was presentable, smart and appropriate for the school environment and mainly black with several flamboyant accessories of jewellery. The young man presented himself with a long ponytail hairstyle. He wore a modern black jacket and trousers with an open necked coloured shirt. He also wore a prominent earring. Both presenters had taken part in the pilot studies on two small groups and felt confident with the format and had contributed to the final test format. Both student presenters projected an image concurrent with one prominent mainstream popular music culture.

The three presenters were all introduced to the groups in the same way:

'Today we have something different. This is Mr./Ms (Name), and they are interested in how you listen to music, so they have a musical game for you.'

The test was introduced using almost the same script as in the previous test. The three presenters gave one slight variation, in that they added the words:

"I really hope that you like the music on the tape, because I have chosen all my favourite music"

In keeping with the suggestion from Radocy (1976), remarks made to the participants "varied in their exact wording in order to avoid a mechanical, unnatural pattern" (p. 122). However, as has been stated previously, all three presenters witnessed the presentation of the other two. Subject matter was kept identical and also the position within the test when it was given.

6.3 Design and procedure

As already stated, the experiment incorporated three conditions **Co** (control), **Em** (experimental male) and **Ef** (experimental female). Participants received the same style sensitivity test already described in experiment one. The test material, introduction and method were used according to the procedures already set down in chapter 4. All participants in all three conditions experienced 20 musical samples of which 10 were classical and 10 were popular examples. Again the order of musical excerpts was as described in the last experiment. Following each of the tests, the participants' papers were checked through. Any papers where the participant had clearly not understood the task; for example where both 'same' and 'different' responses had been circled, were withdrawn. Following this initial removal of papers, the class teacher withdrew an appropriate number of responses at random in order to make the number of responses equal in all conditions.

6.3.1 Schools

The test was carried out in a number of different schools in order to attempt to gain the widest sample of children. The experiment employed a research population of 540 participants; 135 in each age group. In the two elementary age groups, the participants were taken from two different schools, the first school was a large primary school in West Sussex and the second was a large middle school in a South London Borough. The individual schools differed in catchment area; the number and variety of musical activities that were offered by the school and the amount of music and musical activities that the children experienced within the average school week. One of the schools described itself as offering musical appreciation usually relating to music in the school morning assembly. The school presented whole year group

concerts once or twice a year with all children taking part in some form of singing in those concerts. A small recorder group was offered. The second school offered small recorder groups and small group lessons on Violin and Flute / Clarinet were available to selected interested children. Both schools offered some musical activities over and above timetabled music lessons. In year 8 and year 10, the 13/14 years old participants and the 15/16 years old participants came from a large comprehensive school in South London with 6-form entry. All participants followed the national curriculum music syllabus and around 25% of participants received instrumental tuition, took part in the school swing band or school choir. Participants in both year groups and in all three conditions took the test in small mixed class groups. Testing took place at a time usually timetabled for humanities activities in non-class groups. No group was tested in a music, P.E. or games lesson.

6.3.2 Participants

7/8 years old age group.

In this year group, 142 children were used from the two schools. School A had three parallel classes in each year group. One full class group was assigned to one of the three conditions in a random manner. This was achieved according to the order in which the class registers had been delivered to the school office. In school B, the two classes were split into three smaller groups and assigned to one of the three conditions. The Head teacher assigned the three participant groups to the three test areas but had no knowledge of which group would experience which of the three conditions. Final numbers of participants for each condition were as follows; Condition **Co**: 48 participants; Condition **Ef**: 47 participants and Condition **Em**: 48 participants.

10/11 years old age group.

In this age group participants were taken from the same primary schools as the 7/8-year-old participants. School A had the three-form entry, so participants again experienced the test in their full class group. One full class group were assigned to each condition in the same random manner as in the 7/8- year-old age groups. School B again arranged three groups from the two classes and some effort was made to balance male and female participants. The total number of participants tested in year 6

was 144. There were 49 participants in condition **Co**; 47 participants in condition **Ef** and 48 participants in condition **Em**. Responses were further reduced to 135 by random withdrawal by the class teacher. 45 responses were used from each of the three conditions. This made the total number of responses used the same as in the 7/8-year-old age group.

In both 7/8 and 11/12-year-old age groups, the classes or groups in each condition were deemed to be parallel mixed ability by the school having been balanced according to reading ages and SAT results. An additional effort was made to balance the number of male and female participants in each group.

13/14-year-old age group.

In this age group 145 participants were tested. In each of the three conditions participants experienced the test in small groups. Conditions **Co** and **Ef** both had 48 participants and condition **Em** had 49 participants. Following the random removal of papers by class teachers a total number of 135 responses were used, 45 responses from each condition.

15/16-year-old age group.

A total of 139 participants experienced the three conditions. Condition **Co** had 46 participants; condition **Ef** had 47 participants and condition **Em** had 46 participants. In order to keep the number of responses equal across all age groups, the number of test papers was reduced to 135 by teacher random withdrawal. The allocation of participants to conditions was decided according to timetable and availability within the school. Therefore this was deemed to be random as presenters had no influence over which groups were present at the time of the test. The teacher assigning the groups to the conditions had no knowledge of the particular condition participants were being assigned to.

As in the previous experiment, all responses were firstly checked through and all responses were withdrawn in those cases where the participant had clearly not understood the test instructions fully. Again, as in experiment one, a random withdrawal of some papers was carried out by the class teacher in order to balance

total number of responses and equalise male and female participants responses. As in the previous experiment, response sheets were marked with a score of 1 being given were the participant correctly identified the two excerpts as being taken from the same or different piece of music. Where the participant made an incorrect response, no score was given. Each participant had three final scores. Firstly, the total score of correct answers for classical and popular samples, up to a maximum of 20 and two separate scores for the classical and popular up to a maximum of 10 in each category.

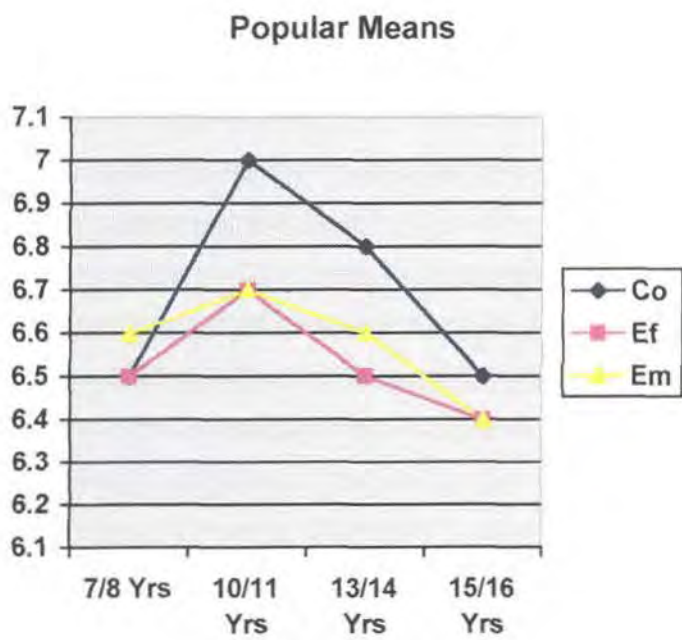
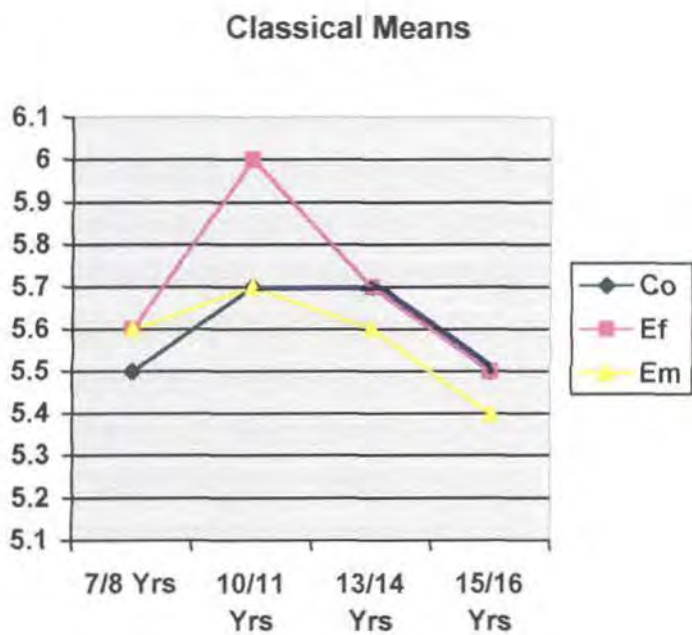
6.4 Results

A 4 way mixed analyses of variance was carried out on the style sensitivity scores: (Musical Style: classical-popular excerpts x gender of participant x age x condition). Further analysis of the classical and popular scores was made by the use of post hoc Tukey HSD tests.

The main effect for age was found to be significant ($F [3,516]=5.47, p=0.001$), Similarly, the main effects for participant gender and Musical style were also found to be significant, ($F [1,516] = 13.31, p=0.00$); ($F [1,516] = 447.85, p= 0.00$). However, the main effect for experimental condition did not reach statistical significance ($F [2,516] = 0.72, p= 0.49$). Only one interaction proved to be significant; that between musical style, condition and participant gender. ($F [2,516] = 4.35, p=0.01$)

The main experimental hypothesis therefore was not proven and the interaction means are plotted in **Fig. 6.1**

Fig. 6.1



6.5 Discussion

The main hypothesis for the experiment was that participants experiencing the style sensitivity test operated by the two younger presenters would possibly perform more accurately than those participants who experienced the test material in the control group. However, the participants did not appear to become more 'open eared' and did not appear to be more tolerant of the listening activity through any form of cultural, kinship or prestige allegiance. The test hypothesis was therefore not proven. There are a number of possible reasons why the hypothesis was not proven. Firstly, there is the possibility that the participants did not consider the change of presenter to be a significant or substantial change. There was therefore no measurable effect on the participant performance. Secondly it is possible that the test material was not susceptible to changes in experimenter gender. Thirdly, there may have been a number of further personality differences or attributes between the three presenters beyond those identified by the experiment (age, gender and cultural style). Further more, it was suggested earlier that the two younger presenters might inspire more accurate listening through the so-called 'prestige' effect. Crozier and Chapman (1981) suggest that there are several reasons why this effect may not work.

-Firstly, the musical stimulus used may be too familiar to participants. That is the style is familiar to them even though participants may be unaware of the individual pieces. Crozier and Chapman found that familiar stimuli were affected less by 'prestige' effects than unfamiliar stimuli.

-Secondly, Crozier and Chapman found that some participants reacted strongly against any form of 'labelling' of musical styles.

-Thirdly, all procedures should be as identical as possible. Crozier and Chapman demonstrated that even slight variations in procedures significantly altered the 'prestige' effect or even eliminated it all together.

-Lastly, they suggested that participant susceptibility to 'prestige effect' also could be affected through the teaching styles participants were most acquainted with.

Although each of the three presenters witnessed each other presenting the test, it is certainly possible that significant differences in personality could have affected the participants' performance. Therefore, any one or a combination of these factors could have prevented the experimental condition reaching a level of significance.

The main effect for age was found to be significant. In common with the previous research, the level of accuracy achieved in the style sensitivity test by even the youngest participants was high with even the youngest participants achieving total scores well above those expected through chance.

Similarly, the main effect for participant gender was found to be also significant. Full data for this can be seen in the appendix. The first study by Gardner (1973a), found a significant difference between the performance of male and female participants, but in this study it was the male participants who performed more often with greater accuracy than the female participants. It is possible that the study by Gardner included an atypical sample of participants and in addition participants were only tested using classical pairs. Alternatively, in this experiment the placing of participants into groups by the schools was primarily based on small friendship groups. It is possible that this produced some form of skew in the population.

Again, the main effect for musical style was found to be significant. As has been seen in all previous research, participants did perform at a significantly higher level on the popular pairs (familiar genre) than on the classical pairs (unfamiliar genre). This finding runs contrary to the experimental hypothesis. It had been suggested that participants in conditions **Ef** and **Em** would become more tolerant of the classical pairs and therefore increase their classical scores to a level more equal to that obtained in the popular pairs.

Classical pairs

The classical scores in all three conditions tended to follow an identical pattern to those achieved in previous experiments with style sensitivity to classical music apparently improving dramatically between the 7/8-year-old participants and the 10/11-year-old participants with the latter age group scoring well above the older two

age groups. Subsequent analysis of the mean scores separately using Tukey HSD tests showed the increase in scores between the 7/8-year-old and 10/11-year-old participants to be significant ($p=0.08$). The Tukey analysis also showed the apparent decrease in scores between the 10/11-year-old participants and the 15/16-year-old participants to be significant. ($p=0.02$).

Although the main effect for experimental condition did not reach a level of significance, there was a significant interaction between the participant gender, presenter gender and musical style. Further examination of the means obtained in the classical pairs using the Tukey HSD tests showed a number of major reactions between participants and the female experimenter. As these reactions appeared to be both positive and negative, depending on age and participant gender, these opposite interactions may have cancelled each other out in the analysis for main effect. The mean scores for male and female participants in all conditions and across all age groups in the classical pairs are set out in Table one (p.169)

Within the 7/8-year-old age group, the male participants in conditions **C** and **Em** appeared to react in an almost identical way regardless of age and culture of presenter. However male participants in this age group appeared to react most strongly towards the younger female presenter (**Ef**) by achieving the highest scores. The female participants in this age group appeared to react most positively towards and scored better with the younger male presenter (**Em**) and seemed to react most negatively and scoring least with the young female presenter (**Ef**).

In the 10/11-year-old age group both male and female participants appeared to react most strongly towards and scored better with the younger female presenter (**Ef**). The males in this age group achieved the lowest scores with the younger male presenter (**Em**) and the female participants scored least in the control group(**Co**)

In the 13/14-year-old age group, the strongest positive/negative reaction occurred in condition **Ef**. The males in this group appeared to score highest in her group but the female participants appeared to react in a negative way and scored very much worse.

Male participants performed worst in the **Em** condition and as already described, female participants scored worst in the **Ef** group.

In the eldest age group, the male participants performed best in the **Ef** condition and performed worst in the **Em** condition. The female participants performed best in conditions **Co** and **Em** and slightly worse in condition **Ef**.

Popular pairs

As in the classical pairs, the main effect for age did reach a level of statistical significance in the popular pairs. Scores in all three conditions also tended to follow the pattern achieved in most of the previous experiments with style sensitivity appearing to improve dramatically between the 7/8-year-old participants and the 10/11-year-old participants with the latter age group scoring well above the older participants. Subsequent analysis of means using a Tukey HSD test showed the increase in scores between the 7/8-year-old and 10/11-year-old participants to be significant ($p=0.07$). The Tukey analysis also showed the apparent decrease in scores between the 10/11-year-old participants and the 15/16-year-old participants to be significant. ($p=0.01$)

Further examination of the means obtained in the popular pairs using Tukey HSD tests separately, again showed a number of interactions between participant gender, presenter gender and musical style. The table of mean scores for male and female participants in each of the three conditions and across all four age groups in the popular pairs are shown in Table one (p.169).

Within the 7/8-year-old age group, the male participants appeared to react strongly in a positive way towards the older male presenter by achieving the most accurate scores. The female participants appeared to react in an opposite negative way to the older male presenter and achieved the lowest scores in the popular pairs. The male participants performed equally with the younger presenters regardless of their gender whereas the female participants performed better with the younger male presenter.

In the 10/11-year-old age group both the male and female participants appeared to react strongly in favour of the older male presenter with the male participants reacting most positively. Again as in the classical pairs, male participants appeared to perform equally well with the two younger presenters regardless of their gender. The female participants in this age group performed equally well with the older male presenter and the younger female presenter. Female participants performance was slightly worse with the younger male presenter.

In the 13/14-year-old age group, the strongest reaction occurred with the older male presenter. Both the male and female participants reacted most positively to the older presenter however the males in this age group appeared to react most positively. Both male and female participants appeared to do equally well with the younger presenters. In the female participants,

In the eldest age group, the male participants again performed better with the older presenter whilst in the other two conditions, male participants performed equal with both younger presenters regardless of presenter gender. The female participants performed better in the **Em** condition and again performed equally well in the **Co** and **Ef** conditions.

6.6 Conclusions

Overall, the experiment would appear to suggest that as a test material, musical examples in a style sensitivity test of the type described and used in this experiment could be a useful measurement of presenter effects. The results suggest some agreement with the conclusion reached by Lutey and Copland (1982) that male and female experimenters often achieve significantly different results from their participants although it is not always possible to predict for any given type of experiment just how participants' responses will be affected by the gender of the examiner and that:

"age and not just sex of the children influenced whether they did better with a male or female tester" (p. 132)

The results suggest that participants' reaction to and their susceptibility to experimenter age, gender and culture can vary as a function of musical style, age of participant and culture/ dress code or behaviour of examiner. Participants appear to respond to different and varying features of the test presenter depending on their age group; younger participants appear to react towards different presenter attributes than do older participants. Younger participants appear to respond in a positive way towards the age and culture of the younger presenters whilst older participants seem to react negatively towards them, perhaps rejecting an almost rival street culture.

" Each fashion that arises is out of date more or less as soon as it becomes widely accepted"

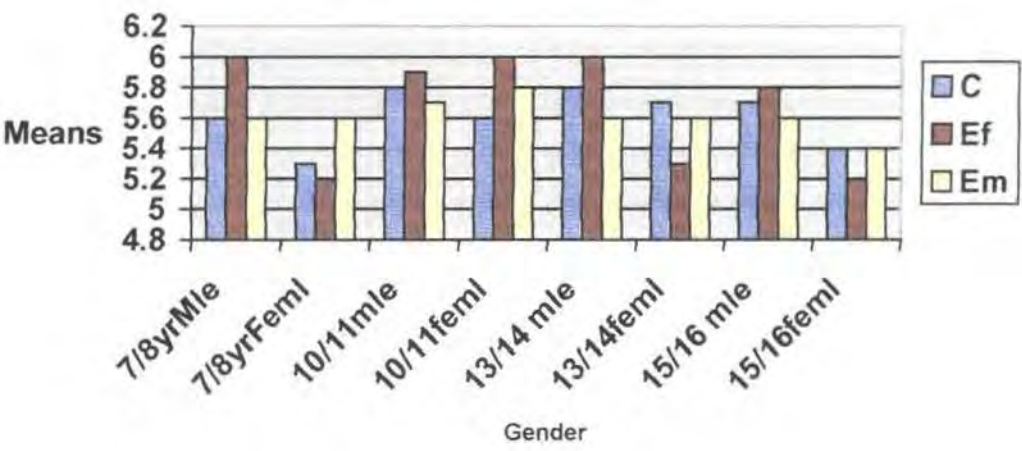
(Hargreaves 1986. p. 191)

It is highly likely that in reality, a wide range of interactions take place between the different age groups and the test presenter. In some instances participants may well identify strongly, either in a positive or negative way with the personality of the presenter. In other instances emotions such as attraction, jealousy or envy may be a source of influence on motivation to perform well in the test as well as a whole range of sub conscious reactions to the individual tester, the musical samples and even the test itself.

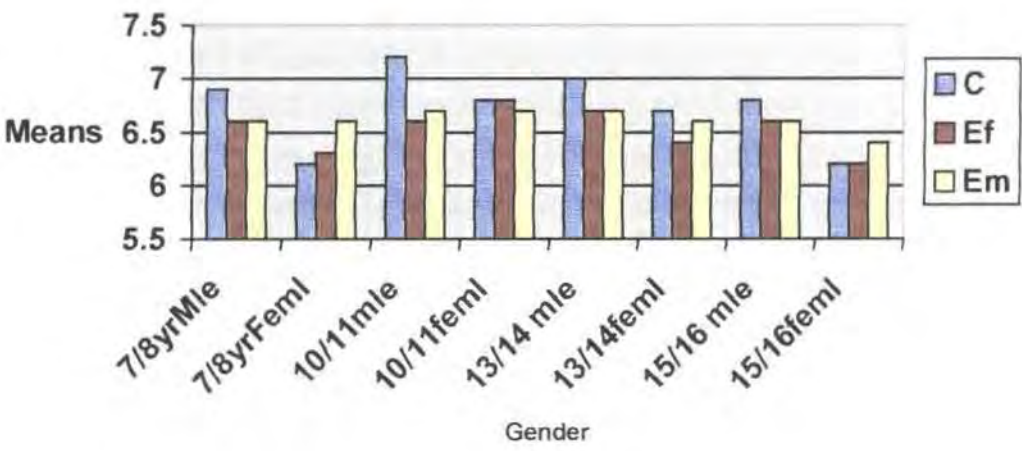
The experiment suggests that strong and interesting interactions do take place as a result of participant / test presenter interactions. Further research projects using a far greater number of and variety of test presenters are required in order to further describe the way in which participants' reactions and listener tolerance change as a feature of participant / presenter age, gender and culture.

Table 6.1

Classical Means



Popular Means



Chapter Seven

Experiment three

National differences in musical style sensitivity: a comparative study in the UK and USA

"Musical development is a major issue in music education. Music educators' actions are based on their ideas about the musical capabilities children possess at various ages and the way these capabilities change"

Koopman, (1995). p49

7.1 Background

Previous research suggests that teenage involvement with music may be similar in all Western cultures (Bjurstrom and Wennhall, 1991) and Cotterell (1996) also suggested that the social and cognitive development of teenagers follows similar lines in most Western societies. The purpose of this comparative study was to explore if pupils from different cultural backgrounds would perform differently on style sensitivity tests. Three groups of pupils aged between 7 and 16 years were used and these pupils were taken from two different Western cultures. One group consisted of pupils from the United Kingdom and two groups consisted of pupils from different regions of the United States. The three groups of pupils all experienced different music education programs. Each of these programs accorded music a different status and each program not only involved different content but also placed emphasis on different aspects of the curriculum. Pupils in each of the three groups also experienced different teenage musical cultures and the local musical cultures were also very different. All groups experienced a different diet of commercial music and the range of commercial music available to them for purchase was also different. The musical styles that were accepted and promoted within the local community were equally varied.

All pupils were taken from the same four age bands, namely 7/8 years old, 10/11 years old, 13/14 years old and 15/16 years old. Group one (**Guk**) employed pupils from two schools in the United Kingdom; one primary school and one secondary

school. Group two employed pupils from two schools on the east coast of the United States (**Gnh**) and Group three incorporated participants from two schools in a mid western state in the United States (**Gi**). The United States was chosen for comparison for a number of reasons. Firstly, the results provide an interesting comparison with Gardner's (1973a) original study. In this study, Gardner suggested that the results from his research might have been affected by the nature of his research population (white, intelligent, middle class Americans from a New England university town). Secondly, the pupils in the American school system experience a very different music education program. American pupils enjoy greater musical resources and music is accorded a much higher status than music education programs in England. Thirdly, it was also considered desirable to keep all the studies within countries in which the primary language was English in order to avoid introducing any uncontrolled influences through language or translation differences. All other aspects of the presentation were kept as close to the original English presentations as it was possible to achieve.

As was noted in an earlier section, although three contrasting schools from contrasting regions were selected for the USA and the UK samples, it is not claimed that any of these schools were in fact representative of the large variation to be found amongst educational establishments in either country. Therefore, it is accepted that without further research and a larger sample of schools, there will be some limitations as to the generalisability of these results.

As a result of the different music education program; the enhanced status of music and the greater resources present within their education system it was hypothesised that American pupils would show greater sensitivity to musical styles than the English pupils. It was further hypothesised that some significant difference would exist between the two American populations as a result of differences that existed in the exposure to and the status of different musical styles within the two regions. Many studies have been carried out on a wide range of participants, both in the U.K. and in the U.S.A, but very few comparative studies have been undertaken. It therefore seemed to be appropriate to explore this issue further.

7.2 National and regional differences

Gardner (1973a) made reference to the overall high standard of performance achieved by participants in his original study. He further remarked that this high level performance may, be attributable to the select and somewhat 'privileged ' research population used in the study. He also concluded that the results could be limited in their generalisability to other populations. The participants used by Gardner were American children from a somewhat privileged area in the vicinity of a top American university. The present experiment explores whether or not any differences occur between the English participants and the American participants given the differences in music education experienced by both populations, and provides comparative data with the original Gardner (1973a) study. Although, as has been reported, a number of different style sensitivity tests have been carried out with other populations, no previous direct comparative study has been carried out between the U.K. and the U.S.A.

Baumann (1960) noted that a difference existed between the musical preferences and everyday listening experiences of East coast and mid West children. East coast children reported experiencing more classical music in the home and also in the output from commercial broadcasting stations than did their mid Western counterparts. Participants from the East coast showed a significant preference for popular music whilst the mid West participants showed a greater liking for the classical examples. Also, the East coast children showed a higher preference for traditional music than did the mid West participants. However, differences were in reality a matter of degree of preference; the results did not indicate that one group accepted music that was rejected by the other. In this sense the results may be a measure of *tolerance* and not necessarily *preference*. Baumann reports that these findings mirror those published in the survey by the American Music Conference (1948), but gave no explanation as to why the differences existed. Some differences were found between 'low and high socio-economic status' participants both in preferences and listening habits.

Baumann's study does supply an interesting perspective on differences between children in Arizona and Maryland. However the study was undertaken in 1960 and

reference is made to ‘the family phonograph’ or the ‘family radio’. Most participants, particularly in the lower socio-economic groups heard music through the ‘juke box’ or through the family music apparatus. Similarly, the amount of music, the variety of styles of popular music and possibly the values attached to music have all changed considerably during the intervening years. The results need to be seen within this context.

Tarrant, North and Hargreaves (2000) found no differences between English and American adolescents reasons for listening to music but did find some differences in listening habits. Adolescent participants from one mid Western state and from one English city were asked to complete a questionnaire about the amount of time they spent listening to music. The questionnaire suggested eleven potential reasons why participants listen to music and participants rated each reason on an eleven-point scale. Participants were further questioned about who accompanied their listening and details of their own individual musical experience was also requested. The results suggested some differences in the listening habits of some of their participants. U.S. pupils reported listening to slightly more music than did U.K. participants. Some interactions were noted between nationality, participant gender, musical experience and listening context, but further analysis of this was not undertaken. There were also interactions between nationality and listening context, with U.S participants suggesting they listened to more music alone or listened alone and with friends in equal measures, than did U.K. participants. Secondly there were interactions between nationality and participant gender:

“U.S males reported listening to music more for self-actualising reasons than did any other group, whilst U.K males reported listening to music less for this reason than did any other group”
(p.170)

Also, some interactions were noted between musical experience and nationality. The authors reported that in their study “it is noticeable that considerably more U.S participants reported high levels of musical experience than did U.K. participants” (p.171). This replicates the findings of this writer and supports the rationale for this present study. The study by Tarrant, North and Hargreaves found that level of musical

experience did influence the main reason why participants stated they listened to music.

A full physical description of the music departments within the schools employed in this study is presented in a later section. However a number of further differences did exist and these are worthy of comment. The description of the status and resources within the American system carries with it the implication that U.K. pupils are less privileged than their American counterparts. Certainly, the funding and the level of resources available to American teachers are much greater. Similarly the status afforded to music, in the schools included in this study, is greater.

Pupils in the U.K. experience a much broader curriculum involving activities designed to develop their capacities in composing, listening, appraising and performing. The 'Programmes of Study' are designed to develop perception and attentiveness in listening activities and pupils are required to discriminate between complex musical elements in a wide range of musical styles. Pupils are made aware of musical elements such as pitch, duration, pace, timbre, texture, dynamics and structure as well as experiencing diverse musical styles such as Duke Ellington, Blues and Folk music alongside Gregorian chant and Stravinsky (D.E.S., 1992). Within the English music education system, pupils should systematically be made aware of the various elements existing within the music of their culture. Using Koopman's (1995) terminology, the English education system systematically assists development within the horizontal dimension with pupils further defining the domain and tracing further dimensions for vertical reconstruction.

The two American states involved in this present study operated music education programs based very much on two prominent musical aspects namely choral singing and instrumental performance, especially 'Marching Band'. Pupils took part in a large amount of practical music, some music appreciation and learned some factual knowledge about composers and learning to read musical notation. Very little or no time was spent on personal or group composition. Therefore subject matter such as form, structure, texture and dynamics were never learned through use in personal

compositions, they were only experienced according to how they had been used by others. Again in Koopman's (1995) terminology, music education concentrated on specific dimensions within the domain. So the two different systems may affect very different aspects of musical development, the U.K. system favouring the horizontal development whilst the U.S.A. system favoured the vertical development.

This difference between the two forms of music education program may reveal some interesting results in a comparative study of style sensitivity. In terms of style sensitivity, this may be one aspect of musical competence that is not affected by formal musical training of either the U.K. or U.S. variety.

7.3 Effects of training

Morrongiello (1992) has reviewed a number of studies into the effects of musical training on children's perception of music. The literature she reviewed suggested that formal musical training (that is, instrumental training), served to enhance the perception of rhythm and tonal properties of melodies. She claimed that,

“ children with training outperform their untrained peers on a variety of musical perception tasks” (p.38)

Morrongiello further noted that the speed with which musical information was processed was greater in those with formal training, and that musical memory was also improved. However, these results were the subject of two limitations. A positive correlation was found between participant's age and the effectiveness of the formal training in increasing participant performance on the perception tasks. As no long-term studies were undertaken, the interactions between age, effectiveness of formal training and participants' subsequent performance were not explored further. Secondly, the research was only carried out on children who had already experienced some musical training prior to the study. No control group of participants without any previous musical training was included. The results may have therefore been influenced by some form of interaction between the previous training and the training within the study itself.

The effects of musical training on the perception of tonal properties of melody have been investigated by a number of authors. Results suggest some degree of variation in the effectiveness of musical training in enhancing the development of individual musical skills; some skills being far more affected than others. For example, training did appear to improve performance on the effective use of tonal structure (Cuddy and Cohen, 1976); the salience of major-minor key changes (Halpern, 1984); and on the discrimination of timbre (Crummer, Hantz, Chuang and Walton, 1988), but training did not appear to lead to improvement in the perception of melodic contour (Bartlett and Dowling, 1982).

Morrongiello, Roes and Donneny (1989) carried out a study on two groups of participants; namely trained and untrained. Those participants classified as 'trained' had received an average of 11.5 months training in violin playing by Suzuki method and the study involved children between the ages of 4.5 and 6 years of age. Participants were presented with a standard 6 note melodic sequence followed by either a repetition of the same melody or by a second melody which was a transformation of the first. The second contrasting melodies were transformed from the original in one of three ways; either the same melodic sequence was used but transposed into a new key; the same melodic contour was used but with minor changes to some intervals or lastly, the total melodic contour of the melodic sequence was changed by re ordering the same 6 notes as the original. The task required participants to identify whether the second melody was the same as the first or a transformation of it. The results suggested that although musically trained participants did perform better overall, performance varied as a function of level of musical training and type of transformation.

The authors therefore concluded that musical training did not assist all aspects of musical perception. Furthermore, this study does suggest that participants as young as 4.5 years are able to both understand and discriminate on a 'same' / 'different' task and that musical training may enhance performance on some, but not all of types of listening task. This feature is explored more fully in a later experiment.

Williams (1972) carried out four studies in which he explored participants' musical preferences and how susceptible those preferences were to change under a number of varying conditions. Using a pre test – posttest format, he explored whether or not musical training and instruction could affect participants' attitudes towards various styles of music. Subsequent experiments explored whether or not the effectiveness of the instruction program in changing participant's attitudes was affected by participant's socio-economic status and / or musical aptitude. Williams used both classical (symphonic, chamber and vocal) and popular (folk and current popular) musical examples with participants recruited from two American colleges. Occupation of the head of the household was used as a determinant of socio-economic status and musical aptitude was measured using Gordon's Musical Aptitude Profile. Socio-economic status and musical aptitude were not found to be significant factors in the effectiveness of instruction in changing musical attitudes. Williams did however note that training which incorporated one style of music and produced a positive change in attitude towards that style of music also tended to produce a "negative shift" in attitude towards other styles of music. He noted :

"instruction in one type of music may alter the attitude toward that type of music as well as some other type" (p.36)

Brittin and Sheldon (1995) carried out an exploration of participants listening preferences and tolerances under two conditions. The first condition explored differences in listener preferences and tolerance between music major and non-music major participants. The second condition explored differences between static and continuous measurements of music listener preferences, using a Continuous Response Digital Interface (CRDI). This recorded participants' continuous preference responses throughout the total playing time of the musical sample. All changes in participants preferences made in response to elements within the music were recorded, and a mean value was calculated. In the static measurement condition, participants were asked to make a once only preference judgement on a 10 point Likert - type scale for each musical sample. The results suggested that participants with more musical experience gave consistently higher ratings in their preferences for classical music than those participants with less musical experience. The style of the testing format also revealed some differences. In the case of music majors, both the static and continuous measurements revealed similar preference means. However, for non-music majors,

the style of testing produced some significant differences. Static preference / tolerance scores were found to be much lower than those obtained in the continuous measurement condition. Gregory (1993) also found significant differences between the preferences for classical music examples between high school participants who took part in musical activity groups and those who did not. In the overall preferences for musical styles, both music and non-music majors disliked the classical, piano only style examples most of all. As this and other previous studies have incorporated classical piano only examples, this may have had some effect on participants' tolerance in previous tests. The issue of musical samples is addressed in a later study.

Geake (1999) explored the performance of musically gifted and non- gifted primary aged children on the various component parts of Gordon's Music Aptitude Profile. (MAP). This involved 10 individual tests divided into three main sections. Tonal Imagery (3 tests including Melody and Harmony); Rhythm Imager (3 tests including Tempo and Meter) and lastly Musical sensitivity (4 tests including Phrasing, Balance and Style). The latter test resembled the style sensitivity test used in the present study with participants hearing musical pairs. They were required to identify whether the pairs were *like* each other (the same), *unlike* each other (different) or say if they were *unsure*.

Geake attempted to explore how musical information was processed and therefore he questioned whether or not identical musical skills were required in all types of musical task. His results suggested that:

“Musically gifted subjects did show superior overall performance on the MAP over normal subjects, most strongly on the tests for tone, but not generally on the on the tests for musical sensitivity” (p.21)

Geake's results suggested that musical sensitivity was a separate construct from musical aptitude and therefore would be least affected by musical training.

7.4 The present study

In summary, the literature reported here has suggested that differences in style sensitivity may occur between participants from different teenage cultures and between participants who experience very different music curricula. The study is based on three main theoretical perspectives. Firstly, teenage involvement with music may be similar in all Western cultures. Similarly, social and cognitive developmental processes may be similar in all children from Western cultures (Bjurstrom and Wennhall, 1991; Cotterell, 1996). However regional and national differences may exist in musical attitudes, preferences and tolerances. These differences may be the result of regional differences in the commercial music diet; dominant attitudes within a region (Baumann, 1960) or listening habits (Torrance, North and Hargreaves, 2000). Secondly, some formal training may affect the process of development in some musical skills (Morrongiello, 1992), whilst other skills are seemingly less affected by training. Thirdly, the accepted status of music in schools and in the community may affect listener tolerance (Brittin and Sheldon, 1995). As previous research (Castell, 1983; Hargreaves and North, 2000) had suggested that social and cultural factors might be affecting listener tolerance and therefore performance in style sensitivity, involving pupils who may accord music more status and be more tolerant of individual styles will be an interesting and valuable exercise.

The purpose of this experiment can therefore be summarised as follows:

- to explore any differences brought about by music training programs, status and musical experience. (Williams, 1972; Morrongiello, 1992; Brittin and Sheldon, 1995; Morrongiello, Roes and Donnery, 1989 and Geake, 1999).

- to further explore any differences in musical attitudes and style sensitivity performance between participants in the American North East and mid-West and the United Kingdom. (Baumann, 1960; Tarrant, North and Hargreaves, 2000).

- to further explore the concern expressed by Gardner (1973) relating to the limitations of his small, middle class and relatively well-educated research population.

The experiment will seek to highlight differences in style sensitivity between three groups of participants from very different geographical areas, cultural backgrounds

and music education programs. The first hypothesis is that participants in group two (**Gnh**) and group three (**Gi**) will achieve more accurate scores in the style sensitivity test than participants in condition one. The second hypothesis is that participants in groups 2 and 3 (**Gnh & Gi**) will display different patterns in their respective mean scores. This may occur for a number of reasons.

-participants in a music education program that places greater emphasis on musical performance will be more able to perceive differences in musical styles.

-participants within an education program that accords music a higher status will be more open eared and tolerant of classical musical samples, and therefore perform more accurately. They will not be subject to the same narrow social preferences imposed by peer pressure.

-participants from different cultural backgrounds with different musical preferences will be more tolerant of all musical styles.

7. 5 Design and procedure

The design of the experiment in terms of numbers of participants, musical samples used and introduction to the style sensitivity test all closely adhered to the design and presentation used in Experiments 1 and 2. The musical samples, recording and reproduction equipment all remained identical to those used in the first two experiments. The test introduction and response sheet were as previously used and described. No test took part during a period designated for musical activity and the writer presented the style sensitivity test in all three experimental groups.

In the next section, we present a brief description of the physical environment of each school music department, and a summary of the salient features that may have contributed towards the perceived status of music in each school. The second section (7.6) details some further information about the musical experiences, listening opportunities and musical culture in which each of the schools were situated. Lastly, a description of the number of the participants in each age group and in each experimental condition is given (section 7.7).

7.6 Schools

A number of differences were apparent in the perceived status of music within the American and English systems as well as substantial differences in the teacher background and the musical resources available. A brief review of these differences is presented here.

The status of music in all the American schools was very high and greater than in the English schools. The following descriptions apply to all the American schools used in the study.

- Music was viewed as a high status and specialist subject requiring a specialist teacher and specialist accommodation within the school.

- The physical characteristics of music departments in all the schools were impressive. In the elementary schools, teachers worked in a large tiered music room with good quality music stands all displaying the school emblem. A large array of good quality percussion instruments were on display. A minimum of six private sound proofed rooms were provided for instrumental teachers. Each school had an additional large hall available for large-scale productions for which priority was given to arts programmes.

In one elementary school in Iowa, the teacher had a radio microphone to assist her voice in the large classroom. The teacher here was also only expected to 'teach'. That is, she was provided with an assistant who had responsibility for the finding, setting out and filing away of all required music for each lesson and also to act as an accompanist in order that the teacher should just 'teach'. The school had purchased and stored a large number of choir 'gowns' for children who took part in school concerts.

- The city education authority provided instrumental teachers, with no costs to the schools. Students were required to purchase an instrument, but special terms were available to help with this.

- Music was a specialist subject and children left their class teacher to come to the music department and take music lessons from the specialist teacher.

-In the high schools, students took part in string groups, choirs or other musical groups all received this tuition during the school day and activities such as choir and orchestra were counted towards their final grades. The school prospectus stated that pupils were encouraged to take these activities in order to become “a fuller and more holistic personality”. Musicians were seen as representing a valuable feature of the school and individual musical tuition or absence from class lessons in order to take part in musical activities, was not seen in a negative way by class teachers.

- American music teachers are trained as specialist teachers and are required to attend continuous postgraduate training on a yearly basis in order to remain fully licensed. Funding is made available for this. Although differing slightly according to the state law, music teachers have to be licensed and trained to take vocal music, instrumental music or elementary music. Some states even require additional training for teaching marching bands. No music teacher is a general music teacher who has responsibility for all areas of music.

-All music teachers met three or four times each year with the local Superintendent of Schools to discuss issues purely relating to music as a subject.

In both American senior schools, a large performance hall was provided just for music. Each hall had tiered orchestral staging and auditorium seating. Again, extra status was awarded to successful music students in the form of ‘gowns’ that were available for those who took part in school productions. The main halls possessed a high standard and wide variety of percussion instruments including timpani and tubular bells.

The English schools reported on in this study, by comparison, operated in small music areas of their usual classrooms. They had some pitched and non-pitched percussion instruments, many of which were not of good quality, neither were they in very good condition. Music teachers were responsible for all aspects of the running of the music curriculum. Instrumental teachers often taught in office space or library areas. In the primary schools the class teacher taught national curriculum music. The specialist music teacher ran any extra musical activities such as recorder groups, choir or small orchestra during lunchtime. In cases of very bad weather this had to be cancelled as

other children needed to use the space. Music appreciation was often taught through music in assembly and 'Composer of the Week' was a popular activity. In this activity, a particular composer was selected (mainly classical). During morning assembly, information on the composer was given and sometimes displays of books and pictures or CD covers were arranged. During the week a selection of music by the composer was played before and after the assembly. Several times each year a popular style of composer was chosen. Recent examples were Paul McCartney, John Williams, Michael Nyman and the Chieftains. In one school the music teacher was part time. Extra musical activities did not count towards any final grade. Status was achieved through performance in school concerts, which took place in the main hall. In both primary and secondary school, both music teachers suggested that some children did not enjoy the stigma of being withdrawn from other mainstream lessons to attend their instrumental tuition.

Some class teachers resented the break in the lesson when pupils had to leave for individual musical instrument lessons and had passed a negative comment on a number of occasions. One music teacher gave an example of a typical negative comment:

“Oh.....do you have to go again!..... as soon as we get on with some work you have to go. You cannot afford to miss this and I can't go over it again just for you. You need to borrow a book make sure you catch up in your own time”

The music teacher made further comment:

“The message is that music is not really work....you know, it's fine to miss music but it's not acceptable to miss this work.... it depresses me sometimes, especially when the music is one thing the child can excel at”

Again it is worth stressing that no claim is being made here that any of the schools used in this experiment were necessarily representative of the UK or US systems. To revert to the terminology of external validity, whilst it is claimed that the sample was representative of the available populations, it is not possible to say that they were fully representative of other populations in different states, regions or schools.

Therefore there may well be limitations as to the extent to which the results may be generalised.

7.7 Participants

The American populations used in this study were taken from two different communities within the United States. One population was drawn from a small town in the mid west state of Iowa, whilst the second came from a small New England town on the east coast, one hour north of Boston, Massachusetts. New England was selected in order to use a population that was close in geographical location to the population used in Gardner's original study. Iowa was chosen because it represented a very different community in the mid west. Many mid western states have large cosmopolitan cities and operate reciprocal teacher licensing arrangements with other states. In contrast, Iowa is very rural with a comparatively small capital city. Communities are small and separated by large distances. European influences are far less evident than in other states in the mid West, resulting in far more isolation.

This isolation was evidenced by a number of facts.

- Few professional people from outside the state moved into the area.

- Current affairs and news programs from outside agencies (such as the BBC), although shown on public broadcast stations in many others mid West states, were not in evidence in Iowa.

- Music stations on local radio presented far more country and western, old time country music or Appalachian music. A smaller number of current popular music stations were available.

- The only classical music station available came through public broadcasting, and this station tended to be dominated by obscure or little known composers. This presented interesting and new opportunities for listeners with an already established interest in classical music. However, it provided an interesting contrast to the population used by Addressi et al. (1995, 1996), who described participants as performing better on classical genres with which they were familiar (such as Mozart).

These authors suggested that participants had learned these styles more thoroughly through their frequent use as soundtracks on television and radio adverts.

Group One (Guk)

In group one (**Guk**) the same style sensitivity test material as employed in experiments one and two was given to English participants from one elementary school and one secondary school. Both schools were located in the south east of England. There were 41 participants in the 7/8-year-old age group and 43 participants in the 10/11-year-old age group. In both of these age groups, participants came from two full classes and they experienced the test in their whole groups. There were 45 participants in the 13/14 year-old age group and a further 42 participants in the 15/16 year-old age group. Each age group consisted of two full classes and participants experienced the test in their whole class groups. A total of 171 participants were tested in this group. Following random withdrawal by the class teacher, as described earlier, a total of 160 responses were used in the final analysis. That is, 40 responses from each of the four age groups.

Group two (Gnh)

In the second group (**Gnh**), participants were taken from two schools in New Hampshire. The town was situated 60 miles north of Boston. Communities were again small and scattered with children travelling up to 10 miles to get to school. Musical tastes were more limited and more similar to the English children than to the mid West American children. Favourite music included heavy metal and current popular bands along with some classical music. Very little mention was made of country and western or any form of folk music although there was a very strong folk dance community present. There was also a reduction in the number of participants who mentioned the church as a musical community and also fewer children took part in musical activities of their own devising. A greater choice of music stations was available but was not fully taken advantage of. Group two consisted of 47 participants in the 7/8-year-old age group and 45 participants in the 10/11-year-old age group. In both age groups, participants again came from two full classes and they experienced the test in those class groups and in their own classroom. Participants in these two age groups again all came from the same school. There were 44 participants in the 13/14 year-old age group and a further 46 participants in the 15/16 year-old age group. As in

conditions one and two, participants came from the same school and each age group consisted of two full class groups. All participants experienced the test within their own class groups and in their own classroom. A total of 182 participants were tested in this group. This was reduced to 160 responses following random withdrawal. 40 responses per age group were used in the final analysis.

Group three (Gi)

In the third group (Gi) participants came from the rural, mid western, community in central Iowa, as already described. A large number of the participants came from a farming background and lived on small isolated farms. Many were third and fourth generation European immigrants. A large number of the participants reported learning two instruments, and a large percentage had experience of choral singing through school and their church choir. Many of the churches offered musical activities; two particular churches offered 'Summer Music Camps'. Many of the pupils took part in the school marching or swing band. Pupils' descriptions of their musical tastes and the music they heard in everyday life suggested a broader range of styles than either the English or East Coast participants. Country and Western music in particular, featured more prominently than amongst the other two populations. Classical preferences were often related to church, for example Bach organ music or choral anthems. However, many of the older participants still favoured Grunge and Metal and younger participants still opted for current bands that were popular amongst young teenagers such as Boyzone and the Spice Girls. The participant's informal descriptions did suggest some support for the Baumann's findings (1960) in that preferences were by degree and acceptance of one style did not indicate rejection of another. Amongst the group 2 and group 1 participants this was not found to be the case. Gnh and Guk participants did appear to more readily reject those styles (e.g. Country and Western and traditional) which were beyond their main preferences.

In the Iowa group there were 43 participants in the 7/8-year-old age group and 43 participants in the 10/11-year-old age group. In both age groups, participants came from two full classes and they experienced the test in those class groups and in their own classroom. Participants in these two age groups again all came from the same school. There were 41 participants in the 13/14 year-old age group and a further 40

participants in the 15/16 year-old age group. As in the other groups, all participants came from the same school and each age group consisted of two full class groups. All participants experienced the test within their own class groups and in their own classroom. A total of 167 participants were tested in this group. Following random withdrawal, a final total of 40 responses per age group was used in the final analysis.

In each case, the writer spent a full day in each school and made an informal assessment of the level of musical involvement and experience of the participants in all three groups. Initially, each music teacher was questioned about the participants' musical experience in school and approximate percentages were given of the number of children taking instrumental lessons and taking part in choral singing. The overall percentage of children taking musical instrument tuition within the whole school was given in each case and other information about staff training, personal development and department budgets were offered. Following each test, almost all participants offered information about their musical activities and their musical experiences, both in school and at home. Three typical responses were:

From Group Gi

Male 11 years. "Sung in school choir since 3rd. Grade and I've sung in church choir. Choir practice... that's one night every week. I have played the trumpet since the 2nd, Grade. I sometimes play the trumpet at church, at Christmas I played it. School Band every week since last year. I borrow CDs from my brother and he has all these heavy metal ones... my mum has got Fantasia and I like that and er.....my gran has got this Mozart, I think it's called , and I like some of that as well"

From Group Gnh

Female 16 years. " Piano, 4 years. Clarinet 7 years. Marching Band. 2 years. School Choir in Elementary School and Church Choir for 6 years. I like the Black Crows, and Bare Naked Ladies and some other stuff"

From Group Guk

Female 13 years. “ When I was in my junior school my friend learned the flute and I wanted to go with her, but only a few could do it and there wasn’t any room.....we used to sing in assembly, and anybody who wanted to be in the choir could sit at the side. I used to sit there but it wasn’t like a real choir because anybody could be in it and we only went so we could sit on the benches”

This anecdotal information appeared to support the fact that the participants within the U.S. music education system were far more able to enjoy the status of music and benefit from access to more musical experiences than their U.K. counterparts. This finding was supported by the findings of Tarrant, North and Hargreaves (2000) who also noted that “ more U.S. participants reported high levels of musical experience than did their U.K. participants. (43 U.S. participants compared with only 8 U.K. participants) p.171.

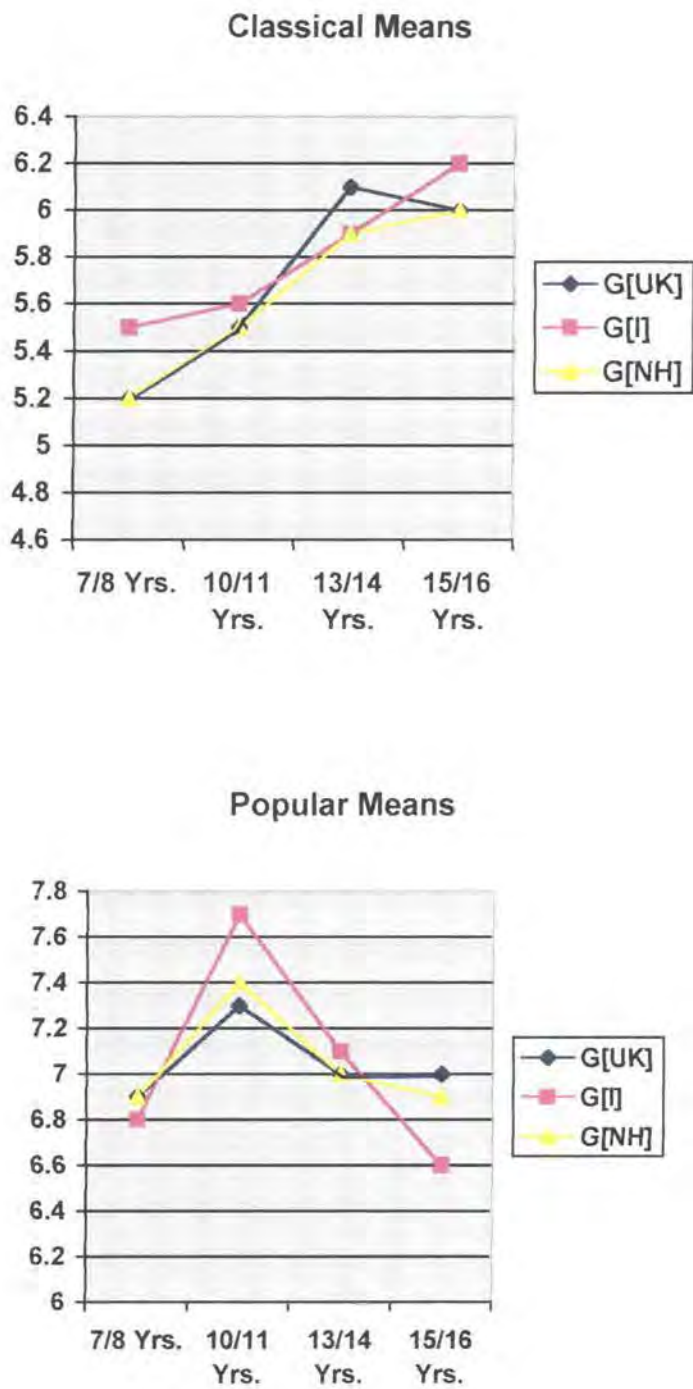
7.8 Results

A three way mixed analyses of variance was carried out on the style sensitivity scores with one within-subjects factor (musical style: classical / popular extracts) and two between subjects factors - age group [4 levels] x condition. As all participants experienced both the classical and the popular pairs, this was a repeated measures analysis. Further analysis of both the individual classical and popular score means was carried out by the use of post hoc Tukey HSD tests.

The main effect for age was statistically significant ($F [3,468]=6.78, p =0.00$); and likewise the main effect for style was also found to be significant ($F[1,468] = 289.82, p=0.001$); whereas that of national / regional group, was not significant. ($F [2,468] = 0.26, p= 0.77$). For the three main two - way interactions, there was found to be no interaction between regional group and age ($F[6,468] = 0.29, p = 0.94$); nor between musical style and regional group ($F [2,468] = 0.23, p = 0.80$). However the interaction between age and style was significant ($F[3,468] = 10.91, p=0.00$). Lastly, the interaction between age, group and style was not found to be significant ($F [6,468] = 0.82, p=0.56$)

The interaction means for the classical and popular scores are plotted in Fig. 7.1.

Figure 7.1



7.9 Discussion

The results of this experiment suggested, in common with the previous studies, that the overall standard of accuracy in both popular and classical pairs was generally high with even the youngest age group achieving scores above those expected through chance. Secondly, in common with Gardner (1973) and Castell's (1983) results, age proved to have a significant effect. 10/11-year-old participants achieved the highest scores in the popular pairs, across all regional/ national groups, and those scores were higher than both the older age groups. In the classical pairs, the 13/14-year-old participants achieved the highest scores in the U.K. group but the 15/16-year-old participants achieved the highest scores in the two U.S. groups.

The main effect for musical style was again significant with participants achieving higher scores in popular musical pairs than in the classical. Therefore the first hypothesis that American participants should perform more accurately was not proven. Similarly, the main effect for regional group was also not significant confirming that the American participants did not perform better than the U.K. group.

The only interaction to reach significance was that of musical style with age. This does suggest that hypothesis two was proven, at least for the popular pairs in that mid Western participants did perform differently from the U.K. and New Hampshire participants.

A further analysis of the classical and popular means separately using Tukey HSD tests confirmed the differences between the means of the four age groups. In the popular pairs, the mean scores of the 10/11 year old participants were significantly higher than that of the 7/8-year-old participants ($p=0.001$). This suggested that most improvement in performance occurred between the ages of 7/8 years and 10/11 years. The mean scores between the 10/11 year olds and the 13/14 year-old age group and the 15/ 16-year-old participants showed a contrasting decrease in performance that also reached statistical significance ($p=0.044$ and $p=0.002$ respectively).

For the classical pairs, the difference in performance between the 7/8-year-old participants and the 13/14 and 15/16-year-old participants was significant ($p= 0.000$) in both cases. Similarly, the increase in mean differences between the 10/11 year-old participants and the 15/16-year-old participants was also significant ($p= 0.003$). These results may add support to the idea that different developmental trends exist for familiar and unfamiliar genres. Within the popular pairs, the 10/11-year-old participants appear to be knowledgeable, attentive and able accurately to identify pertinent features in the musical styles whereas the older participants either become more discriminating and sensitive to more discrete differences within the musical stimulus; or more selective in those styles of music they are willing to tolerate.

The first hypothesis that American participants would perform with greater accuracy than the U.K. participants, was not proven and some comment is required on this. On a social influence level, it is possible that pupils in the American system simply do not realise their enhanced musical opportunities or the enhanced status of the subject and so may remain unaffected. Those aspects of the music departments which impress and raise the status of music in the eyes of fellow adult music teachers from other cultures (e.g. resources, choir gowns, radio microphones and accompanists) may exert little or no influence on participants brought up within the system who simply regard them as nothing extraordinary.

Similarly, regardless of the enhanced status of music within the school, it may simply be the case that musical status imposed (or removed) by peer groups is greater than status imposed (or removed) by school. In terms of the music education programmes, the claim made by Geake (1999) and Morrongiello (1992) may be correct. That is, style sensitivity is one musical skill which is largely unaffected by formal music training, or at least the music education programmes incorporated in the three schools used in this study, but is acquired through more informal / acculturation experiences.

The second experimental hypothesis was proven with respect to the popular pairs. The mean scores obtained did suggest some differences between the three groups. The two American groups did appear to perform in a similar way on the classical pairs but performed very differently from each other on the popular pairs. In the classical pairs, the two American populations experienced music education programs that were most

similar whilst the English participants experienced a very different type of program. The mean scores suggested that the American populations performed in a similar way to each other whilst the English group of participants performed differently.

Similarly, participants from the groups **Guk** and **Gnh** performed in similar ways, whilst the participants in group **Gi** performed very differently. This result may reflect the fact that the English (**Guk**) and East Coast American populations (**Gnh**) also shared similar popular music experiences in terms of the availability of adolescent music and current teenage culture whilst the Iowa participants had a different experience. Certainly with respect to the popular pairs, participants from the Iowa schools appeared to possess less familiarity with popular musical genres than the U.K. or New Hampshire participants. This lack of detailed musical knowledge appeared to influence the performance of the Iowa participants in the popular pairs.

7.10 Conclusion

This experiment suggested that the results obtained by Gardner (1973) were not necessarily dependent upon his use of a selective and un-representative population of participants from a privileged area of New England. In this experiment, similar overall patterns of results were obtained in groups involving English and two different American regions. The results also suggested that in common with the findings of Geake (1999) and Morongiello (1992), style sensitivity requires and operates through a number of different skills that may be largely unaffected by formal music education. The experiment also offers further evidence for the idea that perhaps classical and popular music do have very different developmental trends with each genre affected by very different influences.

Chapter Eight

Experiment Four

The effects of variation of test introduction on style sensitivity

8.1 Background

In experiment four, the effects of a variation in the test introduction were explored. Three conditions were employed. In condition one participants experienced the style sensitivity test with the same “standard” set of instructions as given in the previous experiments one to three. In conditions two and three, the same test content was incorporated as in condition one, but two variations were made to the test instructions. In condition two, the test introduction was changed in order to introduce a competitive element. In condition three, the introduction was changed to incorporate some element of encouragement/ reward. The musical examples in each condition were those used and recorded in the manner already described in experiments one to three. The writer acted as the presenter in each of the three experimental condition and all other procedures and materials were kept as closely as possible to those used in previous experiments. The four age groups namely; 7/8 years old, 10/11 years old, 13/14 years old and 15/16 years old were the same across all three conditions. Hereafter, condition one is referred to as **Co** (control); condition two is referred to as **Cr** (reward) and condition three is referred to as **Cc** (competitive).

Dickstein and Kephart (1972) found increased test performance when participants experienced pre-test motivation. Experimental conditions in which test presenters showed high expectations of participants’ success in forthcoming I.Q tests appeared to produce a corresponding increase in test performance over participants who received no pre test encouragement. Previous studies had concentrated on indirect examiner expectations; that is examiners who tended to behave in an encouraging way towards participants. No attempt had been made to further isolate the different attributes and behaviours which different presenters brought to the test environment. Dickstein and Kephart (1972) operated two conditions in their study. In the first condition, examiners were shown the true grades achieved by participants in previous test situations. In the second condition, examiners were shown participants’ previous

grades that had been altered and made artificially higher than the actual grades had been. The results suggested that examiners who had been shown the artificially high grades appeared to view their participants differently. These examiners held high expectations for the performance of participants in their group and in some way these high expectations were communicated to participants in their groups. As a result of the higher expectations, participants in this condition did perform better. The authors noted that:

" Explicit examiner expectations may be regarded as an attempt to generate a higher level of motivation on the part of the testee" (p.207)

In a second study (Dickstein and Ayers, 1973), a number of college students were given the WAIS intelligence test. In one condition, female participants were given pre test information about the rationale for the test. Participants were informed that the purpose of the test was to explore sex differences in test performance. They were also informed that their scores in the test would be compared with scores obtained from male participants from other high-ranking universities. The female participants were then asked to do their best as they were expected to perform better than the male participants. The results suggested that the female participants in the high motivation condition outperformed other female participants who did not experience the pre-test motivational changes. The authors reported that explicit examiner expectancies could 'dramatically' affect the performance of participants.

Cronbach (1970) suggested that the effects of pre-test motivation could vary amongst different groups of participants.

" Motivation to do a task well or to make a good impression is learned.....The typical middle class child learns to work hard because he obtains praise, tangible rewards, and special opportunities when he achieves well" (p. 61)

The two studies reported here (Dickstein and Kephart,1972) were limited to female college participants and therefore they are limited with respect to gender, age and level of education. Similarly in tests involving verbal skills, no corresponding increase was observed suggesting that either some individual skills are not affected by motivation levels, or that the skill level of a particular population may be stable.

Although the results obtained in the two reported studies are of interest, the note made by Cronbach should also be noted; that the effect of this style of test motivation may be specific to certain types of test material, participant's age, participant's gender and participant's social grouping. However, the authors did suggest that if the examiner wishes to improve participant performance on intelligence tests, then:

"the administration must include a serious effort to arouse maximum motivation" (p. 211)

A follow up study by Dickstein and Ayers (1973) attempted to extend the results of their previous study by further exploring the effects of different motivational manipulations on test performance. Two further attempts were made to extend the generality of the previous study. Firstly, the motivational manipulation was changed from *examiner expectancy* to *incentive*. Participants were urged to perform as well as possible and the five top scores would be rewarded with one dollar. In the second condition, participants were given Raven's Advanced Progressive Matrices test (1962). This test involved a non-verbal performance task and was included in order to establish whether or not motivational manipulation would extend to tests in which there is minimal examiner / participant interaction. The authors suggested that the non-verbal test material included in the APM test, would reduce the degree of examiner-participant-interaction to a minimum. This contrasted sharply with the WAIS test material used in their previous study which involved extensive and constant examiner - participant interaction. The findings of the later study were found to replicate the overall results of the initial study and the authors therefore suggested that the generality of the results could extend to situations involving reduced examiner - participant interaction. The manipulation of incentive was found to affect performance on the WAIS scores but not on the Advanced Progressive Matrices. The authors commented further that the effectiveness of motivational manipulation varied according to the type and level of difficulty of the task required by the test and also according to the amount of experimenter / participant interaction. High motivation appeared to be most effective with test material (WISC) involving constant experimenter / participant verbal interaction. Conversely, high motivation was least effective in tests involving an introduction followed by participants working alone (Advanced Progressive Matrices). The participants for the study were again all female

students from the same college as used in the previous studies, and were classed, by the authors, as being of high intelligence.

Bergen, McManis and Melchert (1971) also carried a study on the effects of pre-test motivation of American fourth grade participants using the WISC Block Design performance. In their study participants experienced tests in three experimental conditions. In the first condition participants were given the standard instructions according to the WISC manual. Participants in the second condition were offered coloured 'chips' for correct responses. This was named as the 'Token condition' and at the end of the test participants could exchange the coloured chips for money. Under the third condition, participants were given verbal reinforcements from a pre-determined list. The verbal reinforcements ranged from 'good', 'fine' and 'very good' for individual responses to 'Very well done. I'll bet you do the next one just as well' at the completion of the task. This was subsequently labelled as the 'Social Reinforcement condition'. The results of their study suggested that pre-test motivation, in the form of a reward, can affect test performance through increasing motivation and that pre-test motivation can affect males and females test performance in different ways. Increased motivation in female participants appeared to cause more accurate performances. This increase in test performance in females occurred with all types of reinforcement, that is in both social and token reinforcement conditions. In contrast the effect of the increased motivation by both social and token reinforcement on male participants was found to be different. Under both reinforcement conditions male participants increased the speed at which they worked on the test, but only increased the accuracy of their test performance under the token reinforcement condition.

Studies by Samuel, Soto, Parks, Ngissah, and Jones (1976) and by Samuel (1977) found that the susceptibility of participants to 'performance expectations' did vary as a function of race, the physical testing environment and the testing materials. In a study by Isenberg and Bass (1974), adult participants who received either verbal or non-verbal reinforcements from the examiner showed consistent, although not significantly higher performance scores on intelligence tests than a control group receiving neither. Non-verbal reinforcement was achieved by nodding of the head, giving participants approving smiles or comments such as "mm-hmm". The verbal-

reinforcement consisted of positive verbal reinforcing statements such as "good", "fine" or "well done". Dorow (1977) explored the attentiveness and music selection / preferences of 76 fourth and fifth grade participants using the 'episodic reinforcement apparatus'. This apparatus allowed participants to select the style of music they wished to listen to and recorded both their choice and the duration of listening to each style. The study also incorporated teacher approval/ disapproval measures. The findings suggested that, " music taught with high approval from the teacher will become more reinforcing, time spent listening will increase" (p.38)

Keogh and Macmillan (1971) explored the effects of motivational and presentation conditions on the performance of participants from different socio-economic, race and intelligence groups. The test material used was digit span tests which were administered to the participants individually. Contrary to the results achieved by Samuel, Soto, Parks, Ngissah, and Jones (1976) and by Samuel (1977), these authors found that test presentation, race and socio-economic background of participants did not appear to cause significant increases in performance on this particular test material. However, they did conclude that motivational conditions did appear to have differential effects on the short-term memory of their participants.

Quay (1971) explored the effects of reinforcement and reward on participants' performance in the Stanford-Binet test. The study was designed to explore the inferior performance of Negro participants in intelligence tests as compared with white participants of similar socio-economic backgrounds. Two factors, namely motivation and language skills were identified as possible sources for further investigation. The study therefore incorporated these two factors into the experimental conditions. In one condition a Negro presenter gave the test instructions in a Negro dialect. In the second condition, a reward of sweets was offered for those who performed well. The aim of the study was to investigate whether changes in test procedures designed to increase motivation would influence participants test scores. The results of the study suggested that none of the experimental conditions facilitated any significant increase in participants' performance on the Stanford-Binet test.

Madsen, Wolfe and Madsen (1969) explored a number of reinforcing effects. They discovered that sixth grade participants' vocal intonation improved significantly over five tuition sessions when they were offered a monetary reward. However, they also noted that participants who did not receive the reward also improved significantly over the same period and the authors concluded that the musical task itself acted as its own reward. Greer, Randall and Timberlake (1971) attempted to explore this reinforcing effect further by using two contrasting styles of music. This was a relevant study in the context of the present experiment in that it again explored differences between rock music and more traditional classroom music. Testing consisted of vocal pitch acuity of participants whilst singing different ascending intervals and conditions involved participants receiving different rewards according to the level and duration of attentiveness they showed. One condition offered a reward of listening to rock music; the second condition offered the reward of general classroom music listening; the third condition offered a monetary reward whilst the fourth condition offered no reward. The results suggested that although the influence of the pattern of reward on participants' performance was different over time, all participants in the reward / reinforcement conditions made significant improvements over those who were offered no reward. This study therefore suggests that participants in the two conditions involving musical reward and increased motivation due to competition, should perform significantly better than those participants receiving neither the offer of a reward or with increased motivation.

8.1.1 Summary

In summary, previous studies have explored the possible effects of a variety of pre test motivational manipulations using a variety of test materials on a broad range of age groups. Studies using pre test high expectations, verbal praise and token rewards on adult participants did find significant increases in test performance (Dickstein and Kephart, 1971; Dickstein and Ayers, 1973 and Bergen, McManis and Melchert, 1971). Whereas Quay (1971) found that neither using a reward of sweets nor making changes to the introduction appeared to influence participants' test performance. Dorow (1977) did find teacher approval to be significant in reinforcing music listening behaviour in participants and Greer, Randall and Timberlake (1971) found music listening had a natural reinforcing effect on participants' behaviour in musical settings.

Previous research therefore does suggest that some further exploration of pre test motivation, particularly with musical test material, could produce some interesting effects and interactions. In the present experiment, one verbal reward (praise) condition and one increased motivation condition have been selected. Some thought was given to the way in which motivation and reward could be introduced into the change of test introduction. With reference to the reward condition, the school was unhappy with any form of monetary reward being given to one group of children. Secondly, the possibility of non-verbal behaviour and presenter manner increasing motivation or conveying increased expectations was noted earlier (Dickstein and Kephart, 1972). It was therefore felt that simply adding verbal or non-verbal rewards to one group may not be sufficiently different from the control group. Thirdly, pre test motivation has been previously found to be least effective in test situations involving a format in which there is a test introduction followed by participants having to work on their own; as in this style sensitivity test (Dickstein and Ayers, 1973).

It was therefore decided that the music listening reward might be most appropriate. As the test was directly related to musical behaviour, a musical reward appeared to be a natural extension. It was also acceptable to the school as being least disruptive; as opposed to extra playtime or a similar reward. Also it had been noted earlier that musical listening might be a self-reinforcing activity and the studies by Madsen, Wolfe and Madsen (1969) and Greer, Randall and Timberlake (1971) had found it to be a suitable reward.

In the second instance, a competitive element was chosen. As Cronbach (1970) had pointed out, the susceptibility to increased motivation through praise or reward; that is the desire to do well and please may be a 'learned behaviour' and therefore it might be more prominent amongst participants of certain socio-economic groups. Therefore in order to utilise an alternative method of motivation manipulation, a similar competitive approach to that used by Dickstein and Ayers (1973) was adopted.

The hypothesis for the experiment is that participants experiencing the style sensitivity test under conditions two and three will perform more accurately than those in condition one due to the increase in motivation. The experiment seeks to

explore two further points. Firstly, if and how musical style sensitivity can be affected by increased motivation or reward and secondly, if participants in all age groups are equally affected.

8.2 Procedure

In condition one, participants experienced an identical style sensitivity test to those used in experiments 1 to 3. All procedures were kept as similar as possible to those previous studies. This group was designated as a control group and subsequently labelled Condition C.

In condition two, participants experienced the standard style sensitivity as in the control condition but with no change to the procedure. Prior to the introduction being read, participants were given the following information:

Before you start, I would like to say thank you for taking part in this exercise today. You will shortly be asked to take part in a musical activity that involves listening to some short pieces of music. You may or may not enjoy all the music you hear but I would ask that you listen carefully to each of the pieces. I am pleased to say that at the end of the exercise, as a reward, you will be able to choose two of the pieces to listen to all the way through. It's like a "thank you" for taking part.

The test was then administered in an identical way to Condition C. Following the testing procedure, participants voted on the pieces they wished to hear all through. Condition two is subsequently labelled as Condition Cr (reward).

In condition three, the test was administered in an almost identical way to the control group. However, following the introduction, which was read as standard, the writer put down the instruction sheet and added the following:

You have just heard the instructions that you are asked to follow. Now, I should not tell you this, but I am going to give you some information about these musical experiences. Some time ago, this experiment was carried out in England and in America. What they discovered was this; the American children always scored higher results than the English children. The person who carried out this test was an American and he said it was because the American children (the word 'subject' was used for older age groups) could listen better and that the English children would only really listen to music

they really liked themselves. So today, could you please listen really hard whether or not you like the music and really try the best you can because when this is repeated again in America next month, we don't really want to see the English children coming second again. So thank you very much for really paying attention and listening and I am sure you will all do well.

All other details of the procedure were kept as close to the control condition as possible. This condition is subsequently labelled as Cc (competitive).

8.3 Schools

The experiment employed a sample of 360 participants; 90 in each of four age groups. In the two elementary age groups, the participants were taken from one large middle school in West Sussex. The school described itself as being 'not particularly strong on music' but did offer some musical activity over and above timetabled music lessons. The school had a designated teacher with responsibility for music and offered National Curriculum music to a standard rated as 'satisfactory' by Ofsted. All participants completed the test in the school library.

In year 8 and year 10, the 13/14-year-old participants and the 15/16 years old participants came from a large comprehensive school in South London with 6-form entry. All participants followed the National Curriculum music syllabus and around 25% of participants received instrumental tuition. Participants were tested in small class sized groups of approximately 17 pupils. Testing took place during the afternoon at a time usually timetabled for humanities activities. Testing did not conflict with any other musical activity.

8.4 Participants

7/8-year-old group.

In this year group, 99 children were used and the testing was carried out during an afternoon period in mixed class groups. The school had three parallel mixed ability classes in each year group. Each of the three class teachers divided the children into three groups of approximately ten children of mixed ability. Each group of ten children was eventually assigned to one experimental condition. The total number of participants in each experimental condition therefore consisted of approximately ten participants from each of the three classes. Children were divided initially by

friendship groups with further adjustments made to balance gender and ability. Within each class the groups of ten children were designated by the name of one pupil in the group. The order in which the groups attended the test session was made by tossing a coin. The decision as to which group experienced a particular experimental condition had been made by one class teacher withdrawing tokens from a bag. The three identical tokens were either marked with a **Co**, **Cr** and **Cc**. The conditions were presented in the order the tokens were withdrawn from the bag. The final numbers of participants for each condition were as follows; Condition **Co**: 32 participants; Condition **Cr**: 34 participants and Condition **Cc**: 33 participants. As described earlier and in common with the previous experiments 1 to 3, the number of responses was reduced to an equal number of 30 for each of the three experimental conditions.

10/11 year old age group.

Participants in this age group were taken from the same primary school as the 7/8-year-old participants. A total of 97 participants were tested. The three class groups were each split into three groups of approximately ten children in the same manner as already described for the 7/8 year-old children. The assignment of groups to experimental condition was also carried out in a similar manner to that employed in the younger age group. The total number of participants in each condition was as follows: Condition **Co**: 33 participants; **Cr**: 32 participants and **Cc**: 32 participants. Responses were again reduced to an equal number of 30 per condition.

In both the primary age groups, the classes or groups in each condition were deemed to be parallel mixed ability by the school having been balanced according to reading ages and SAT results. An additional effort was made to balance the number of male and female participants in each group.

13/14-year-old age group.

In this age group 99 participants were tested. In each of the three conditions participants experienced the test in small groups. 33 participants were present in each of the three experimental conditions. The participants were of mixed ability, and represented a small number from each of the six class groups. Participants were already divided for humanities groups and came to the test according to the school

timetable for this year group. The decision as to which condition participants received was made by token selection as already described in the previous section.

All conditions had 33 participants and responses were withdrawn in the manner already described. 30 response sheets were used in the final data analysis.

15/16-year-old age group.

A total of 94 participants experienced the three conditions in this age group. The final numbers in each condition were as follows; Condition **Co**; 31 participants; condition **Cr**; 31 participants and condition **Cc**: 32 participants. Selection of participants, conditions and withdrawal of papers were done in an identical manner to that already described in the 13/14 year-old age group.

As in experiments one to three reported earlier, response sheets were marked with a score of 1 when the participant correctly identified the two excerpts as being taken from the same or from different pieces of music. Where the participant made an incorrect response, no score was given. Each participant had three final scores comprising the total score of correct answers for classical and popular samples, up to a maximum of 20 and two separate scores for the classical and popular up to a maximum of 10 in each category.

8.5 Results

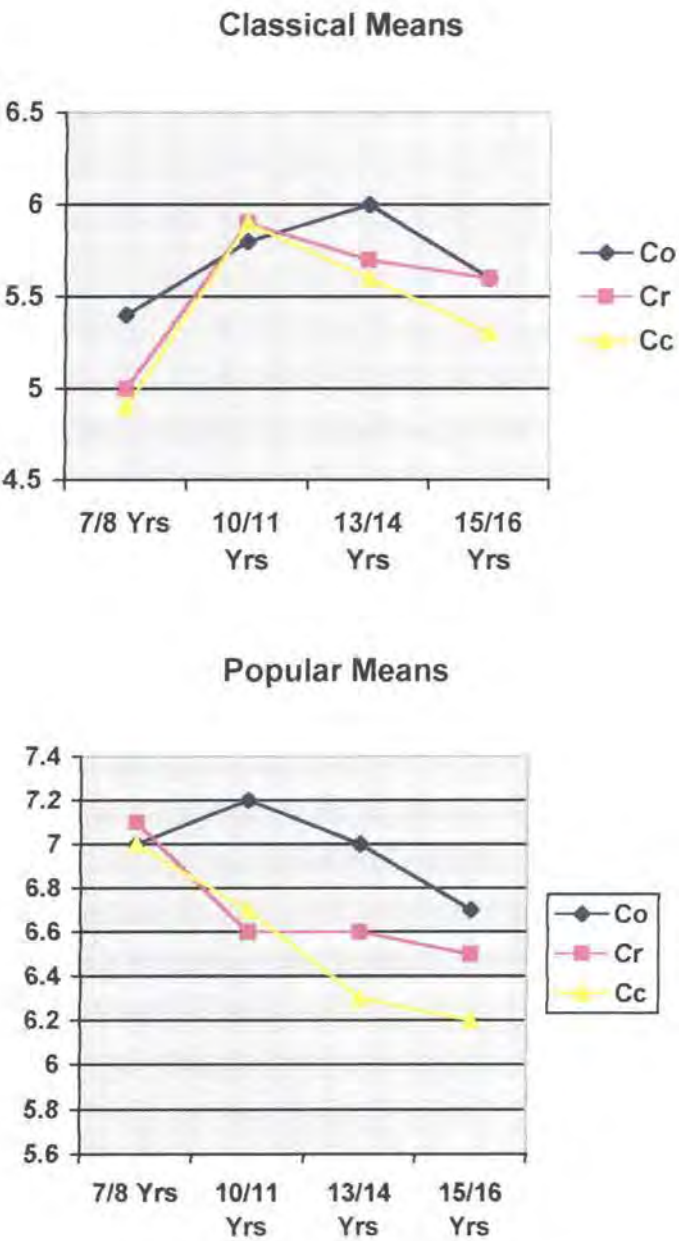
A three way mixed analyses of variance was carried out on the style sensitivity scores with one within-subjects factor (musical style: classical / popular extracts) and two between subjects factors - age group [4 levels] x condition. As all participants experienced both the classical and the popular pairs, this was a repeated measures analysis. Further analysis of both the individual classical and popular score means was carried out by the use of post hoc Tukey HSD tests.

The main effect for age was statistically significant ($F [3,348] = 3.09, p = 0.03$); and likewise the main effect for style was also found to be significant ($F [1,348] = 224.94, p = 0.00$); whereas that of experimental condition, was significant. ($F [2,348] = 3.84,$

$p=0.02$). For the three main two - way interactions, there was found to be no interaction between experimental condition and age ($F[6,348] = 0.31, p = 0.93$); nor between musical style and condition ($F [2,348] = 0.59, p = 0.56$). However the interaction between age and style was significant ($F[3,348] = 8.23, p=0.00$).

Lastly, the interaction between age, group and style was not found to be significant ($F [6,348] = 0.78,p=0.58$)

The interaction means are plotted in Figure four.



8.6 Discussion

The main hypothesis was that participants experiencing the style sensitivity test in the two conditions, involving pre-test motivation through praise or through the introduction of a competitive element, should perform more accurately than those participants who experienced the test material in the control group. However, this hypothesis was not proven. There was found to be a significant effect due to experimental condition but this was a negative shift; that is, participants in these two conditions achieved lower mean scores than participants in the control condition.

Secondly, in common with previous research, the level of accuracy achieved in the style sensitivity test by even the youngest participants was high, with most participants achieving total scores above those expected through chance.

The main effect for age was again found to be significant with 13 /14 year old participants performing best in the classical pairs and 10/11 year old participants performing best in the popular pairs. The main effect for musical style was again found to be significant with participants performing at a significantly higher level on the popular pairs (familiar genres) than on the classical pairs (unfamiliar genres).

The only interaction to reach significance was that of musical style with participants' age. Further analysis again suggested that different developmental trends seem to exist for classical and popular musical styles.

Classical Pairs

As in the previous experiment and in common with the findings of Gardner (1973a) and Castell (1983), the main effect for age was significant for the classical pairs. Although this finding was the same as experiments one to three, the pattern in the participants' performance did vary. In the control condition, participants obtained results similar to those obtained in the previous experiments one to three, in that accuracy appeared to increase between the ages of 7/8 years old and 10/11 years old. Subsequent further analysis of the results using a Tukey HSD test showed this increase to be significant. ($p=.00$). A further increase in accuracy appeared to take

place between the 10/11 year-old and 13/14 year-old ($p=.005$) and a decrease in apparent accuracy occurred between 13/14 year-old and 15/16 year-old participants, though this decrease was not statistically significant. Participants in the 13/14-year-old age group obtained the highest scores.

In Conditions Cr and Cc, the overall patterns of scores were similar with style sensitivity appearing to increase between ages 7/8 and 10/11. Participants in both conditions then showed a decrease in style sensitivity between ages 10/11 years old and 15/16 years old. In condition Cr, participants between ages 13/14 years and 15/16 years appeared to perform almost identically. The significant feature of these scores was that participants in the reward condition appeared to perform less accurately than the control group whilst participants in the competitive group appeared to perform less accurately than participants in both the other two conditions, though none of these reached statistical significance (C / Cr: $p=0.95$; C/Cc: $p=0.24$; Cr/Cc: $p=0.39$)

Popular Pairs

As for the classical pairs, the main effect for age did reach statistical significance. The mean scores gained by participants in the control condition followed a similar pattern to those obtained by participants in the previous three experiments. Participants showed an increase in performance between ages 7/8 years and 10/11 years followed by an apparent decrease in performance between 10/11 year old and 13/14 year old and a further decrease between the 13/14 year-old and the 15/16 year old participants. Mean scores in the control group were found to be significantly different between ages 7/8 years-old and 15/16 years old. ($p<0.002$).

In the Cr condition the participants' mean scores displayed a different pattern from the previous studies with participants in the 7/8-year-old age groups achieving the highest scores. Participants' scores then appeared to decrease with participants in the 10/11-year age group, 13/14-year age group and 15/16-year age group apparently achieving almost identical scores. In the condition Cc, the scores again appeared to follow a new pattern. Participants in the 7/8-year-old age group appeared to score most accurately. Participants' mean scores then appeared to decrease continually across the three subsequent age groups. Yet again, as in the classical pairs, the apparent performance of participants in the reward and competitive conditions

appeared to be well below the scores achieved by those participants in the control condition.

Taken overall the results suggest that style sensitivity tests, at least as operationalised here, can be subject to the effects of motivational manipulations. However, as reported in earlier research, susceptibility to these manipulations can vary as a function of participant age, type of manipulation and test material. In both the classical and the popular pairs, younger participants in the 7/8 and 10/11 year age groups, appeared to be equally susceptible to both reward and competitive conditions. 13/14-year-old participants became more susceptible to the competitive condition than to the reward condition. This may have been because they disliked all the music they could be rewarded with, preferring instead to listen to their own music during the break. The oldest participants appeared to be unaffected by the reward condition but did appear to react to the competitive condition. Again, this could be accounted for by the lack of prestige the type of reward had amongst these participants.

The main finding of the experiment is that participants in the two reward / competitive conditions appeared to perform with less accuracy than participants in the control group for both the classical and the popular pairs. This finding runs contrary to the experimental hypothesis, and requires some comment. A possible explanation may come in the form of anecdotal evidence. In all three previous experiments, there appeared to be a significant difference in the response time required by the younger and older participants. The writer noticed that younger participants appeared to record their response within a very short space of time, often within a second of hearing the opening music from the second musical pair. Correspondingly, older children frequently took much longer over their decision and a review of the answer sheets showed that older children more frequently demonstrated some uncertainty over the answer. The evidence for this was that older participants more frequently circled both 'Same' and 'Different' responses. One response was then crossed out and the actual word 'Same' or 'different' was written in to denote their response. In two instances this occurred up to four times.

This evidence is obviously anecdotal since no record was made of the participants' response times. Similarly, participants' response times did not feature in the design of this experiment but is explored in a later experiment. It appeared that younger participants in the two reward / competitive conditions took much longer to respond to the musical pairs than participants of the same age in both the control condition and in the previous experiments.

One related issue to this would be as to whether or not any significant difference occurred as a result of gender. Males for example, are frequently more competitive than females. However, as reported earlier, Bergen, McManis and Melchert (1971) discovered that the effects of increased competition / motivation can be demonstrated differently in male and female subjects. These authors reported that whilst females responded in an identical way to increased motivation and reward under several conditions, any difference in male performance varied as a function of the experimental condition. That is increased performance in males could be expressed in terms of performing the task more quickly or as increasing their accuracy on the task. The exact nature of the increase in performance was related to the type of reward used in the condition. Therefore within the present study, gender related issues were not explored further. This was due to the fact that if males were to view their time taken on the task as the way in which to increase their performance, the instrument of measurement was not designed to highlight this as an effect. Gender effect was therefore not selected as an independent variable.

Participants in the two reward / competitive conditions did respond to the change in introduction, and to the increase in motivation through either reward or competition. Participants in the younger age groups seemed to show a tendency to respond quickly to musical pairs as they hear one or two prominent differences within the music that are unrelated to style (e.g. volume, production, tempo). In contrast, older participants become more aware of stylistic differences; that one piece of music may in fact contain a number of different styles and changes in instrumentation and possibly tempo and they consequently make more cautious and considered responses. The effect of increase in motivation appears to have caused the younger participants to take more time and make more considered decisions.

Younger participants lacked the musical knowledge of older participants. Therefore, the greater consideration of the music brought about by the greater wish to perform well seems to have created a decrease in performance and not an increase. Within this form of style sensitivity test, younger participants operate musical discrimination through a simple ‘matching’ process, concentrating on major, prominent artefactual elements within the music. In contrast to this, older participants operate a more sophisticated process that incorporates musical knowledge and aural discrimination.

The results of this experiment suggest that within the classical pairs, through the modifications in the test introduction, younger children are motivated to move away from their rapid style of discrimination in which they match prominent features of the music and towards a more sophisticated process for which they do not possess the appropriate musical knowledge. Hence their mean scores are much lower. Within the popular pairs, i.e. with more familiar genre, the increase in motivation, and greater ‘familiarity’ with the musical style produce more accurate performance. This combination reaches its optimum performance level at age 10/11 years.

8.7 Conclusion

This experiment suggests that participants in different age groups carry out the style sensitivity in different ways. Younger participants appear to focus almost immediately on the total sound and some prominent feature of the musical sample e.g. rhythm, beat, instrumentation. Older participants appear to apply their greater musical knowledge to the multiple samplings they take of the musical excerpts; often throughout the entire duration the excerpt is playing.

The study also raises two questions which indicate a possible fruitful area for further research. Firstly, can the anecdotal evidence for younger participants making more rapid responses be confirmed empirically and secondly, if young participants can be motivated to make more considered and prolonged responses, what change in the discrimination process occurs and is it similar to the more prolonged responses made by older participants. Some effort is made to address these issues at a later point in the study.

8.8 Summary

The previous four studies carried out and reported in this thesis have attempted to explore a number of variations to the test methodology which may in some way have contributed to the difference in performance by participants in tests of style sensitivity, as operationalised here. In sum, the comparative results from the four studies are presented below.

Experimental Condition	Highest Mean	Age Effect	Condition Effect	Style Effect	Further Comment
Exp. 1					
Music	Class. 13/14 yrs. Pop. 10/11 yrs.				S S appeared to decrease with age
Non Music	Class. 15/16 yrs. Pop. 10/11 yrs.				
		Significant	Not significant	Significant	
Exp. 2					
Presenter					
Younger Female	Class. 10/11 yrs. Pop. 10/11 yrs.				
Younger Male	Class. 10/11 yrs. Pop. 10/11 yrs.				S S appeared to decrease with age
Older Male	Class. 10/11 & 13/14 yrs. Pop. 10/11 yrs.				Presenter gender also appeared to be significant
		Significant	Significant	Significant	
Exp.3					
UK	Class. 13/14 yrs. Pop. 10/11 yrs.				
USA (NH)	Class. 15/16 yrs. Pop. 10/11 yrs.				S S appeared to decrease with age in UK
USA (Iowa)	Class. 15/16 yrs. Pop. 10/11 yrs.				S S appeared to decrease with age in USA (class.)
		Significant	Significant	Significant	
Exp.4					
Reward	Class. 10/11 yrs. Pop. 7/8 yrs.				S S appeared to decrease with age
Competition	Class. 10/11 yrs. Pop. 7/8 yrs.				
		Significant	Significant	Significant	

Chapter Nine

Experiment five

Variation in contrast between stimulus pairs: the effects of style and chronological period.

“The sublime genius of one generation is the kitsch of another, while supposedly minor composers are taken up as causes by performers and critics of a later era, who resurrect their compositions and elevate them to the pantheon. If that holds true for the music of the past, how much more must it be the case with the music of one’s own time. It is more readily apparent in the field of pop music than in classical music”

(James. 1995. p.229)

In this study three conditions are employed. The first control condition (Co- control) adopts the same style sensitivity task already used in experiments one to four. The second condition (Cc - contrast), involves a second test in which the musical material is changed. This condition employs a similar test procedure to the format used in the first four experiments but incorporates a change to the musical excerpts used. The excerpts are taken from the same four classical periods but an attempt is made to equalise the level of stylistic divergence between the classical and popular pairs. In the third condition, the same experimental procedures were again followed as used in experiments one to four, but the musical material was again manipulated. In this condition (Ce-eras), the change in the level of stylistic divergence was achieved by extending the chronological time period between the four musical eras used.

9.1 Background

As reported earlier, results from previous studies into style sensitivity have consistently found differences in participants’ performance in the popular and classical pairs. In addition, the previous studies carried out by Gardner (1973a) and by Addessi et al. (1995) found contrasting results between classical music excerpts taken from close and from distant chronological periods. This experiment attempts to explore these issues further by exploring one possible cause of the apparent difference in participant performance in the classical and popular pairs.

There are a number of possible reasons as to why participants have consistently performed with more apparent accuracy in the popular excerpts.

- 1) Participants may be processing the musical information in the popular and classical pairs in a different way. This may be due to the different complexity of musical elements inherent within each musical style; the participants knowledge base of each musical style ('novice' or 'expert') or a combination of the two (Gardner, 1973a, Koroscik, 1997).
- 2) Participants' music education programme or the prominent musical vocabulary employed by their immediate peer group (Brittin and Sheldon, 1995; Flowers, 1983, 1984, 1988), may direct their attention towards particular musical cues. These cues may be more prominent in one style of music than the other.
- 3) Differences may exist between the musical cues present in the classical and the popular pairs. These differences may impose a difference in the levels of analysis used in processing the classical and popular pairs (Nattiez, 1990; Meyer, 1989).
- 4) Participants' may be influenced by social or peer group pressure which causes them to dismiss certain musical styles (Castell, 1983).

The background research for the experiment is outlined in three sections and provides a context for the four points listed above. Section one outlines a number of further issues arising from the work on style sensitivity by Gardner (1973a), Addessi et al. (1995) and Castell (1982, 1983). This research is pertinent to points one, three and four listed above. The second section details some pertinent literature on the interaction between musical test material and participants' musical knowledge (Brittin and Sheldon, 1995; Flowers, 1983, 1984, 1988). This may provide a number of possible explanations for the difference in performance found in both the experiments in this thesis and in previous studies. This literature addresses the second point detailed above

The third section outlines two models of stylistic analysis, the first by Nattiez (1990) and the second by Meyer (1989). This section includes details as to how participants may perceive artistic stimuli on a number of different levels. These writers argue that the level at which any analysis of a musical work occurs can vary as a function of the immediate purpose of the analysis and the individuals' knowledge base. The last section links together a number of strands from the previous research with the main rationale for the study. This addresses the third point given above.

9.1.1 Previous empirical work

Gardner, 1973a & Addessi et al. (1995)

Two of the earlier studies into the development of style sensitivity (Gardner, 1973a; Addessi et al., 1995) provided two interesting and contrasting results. Gardner (1973a) noted that his participants appeared to be more sensitive to musical eras that were further apart chronologically, e.g. Greater apparent sensitivity to Baroque and Romantic rather than Baroque and Classical. In contrast to this, Addessi et al. (1995) found no such effect and argued that at least within the classical pairs, participants performed better on those classical styles with which they were most familiar (e.g. Mozart, Baroque), and that stylistic sensitivity varied not as a function of the chronological distance between musical periods, *but as a function of stylistic divergence from those musical systems (patternings) most familiar to and already internalised by the participant*. That is, participant's sensitivity varied as a function of the distance a particular excerpt diverged from those styles which were most familiar to them. They suggested that participants had a greater familiarity with some musical styles, probably due to their popularity and persistent use in the media. This in turn, they argued, allowed participants to more fully internalise the 'patternings' which signified that style. This argument put forward by Addessi et al. (1995) raises an important issue and requires further comment.

If the thesis set out by Addessi is true, then the degree of style sensitivity in children can be explained in terms of their exposure to musical styles. The greater the exposure to an individual style; then the greater the opportunity the participant has to internalise the 'stylistic patternings' of that style. In their study, carried out on Italian children, the authors discovered that their participants performed better on excerpts from the

Baroque period and by Mozart. It was argued that the predominant use of these styles of music in the media promoted participants' knowledge of them. This raises a number of questions.

Firstly, the relationship between musical education experiences and musical experiences from the media needs to be further explored. Did the music heard in the media by the Italian participants augment the musical experience in schools; or did it override their school musical experiences. What is the exact relationship between these two forms of listening experience? If children 'internalise stylistic patternings far more efficiently from the media than from their music education, this carries a number of ramifications for music education programmes. Secondly, if indeed participants had internalised patternings from the music of Mozart and the Baroque period, what was the reason for this? Did participants hear far more of these styles of music or alternatively, did participants hear all styles of music equally and yet some intrinsic component or feature of these particular musics ensure their patternings were internalised first?

A further issue outlined by Tafuri et al. (1994), was that "controlling and isolating different, stylistically - pertinent musical traits" in the test excerpts was difficult. They suggest the difficulty arises because classical styles are more complex, and the number of possible variations to musical elements is much greater than in popular musical examples. Additionally, Tafuri et al. (1994) argue that as so little is known about the mechanisms by which people internalise stylistic traits, it is therefore difficult to know which musical elements to isolate and control. As a result of this, even under conditions where some musical traits have been isolated and controlled, the level of stylistic divergence amongst the classical excerpts and the level of stylistic divergence amongst the popular excerpts can still be very different.

If participants' cognitive processes in the style sensitivity test involve identifying, discriminating and matching the most prominent cues or artefactual elements within the music, there appeared to be many more such prominent cues amongst the popular pairs than the classical pairs. That is, it seemed easier to discern the difference

between a piece of ‘Thrash’ and a piece of ‘blues’ than it was to discern the difference between a piano piece by Chopin and a piano piece by Rachmaninoff. In the studies by Addessi et al.(1995) and Hargreaves and North, (1999) the classical pairs were specially selected to be ‘more homogenous’ and varied only in style whereas the popular pairs differed in many other ways. Therefore, in the classical pairs participants may need to identify ‘patternings’ in order to arrive at the ‘correct’ answer. However, in the popular pairs they may have to simply match a variety of more prominent musical cues.

Castell (1983)

As reported earlier, Castell (1983) found that participants performed more accurately on popular than on the classical pairs and the 9 year-old age group performed better than the older 11 year-old age group. Castell found that the ‘poorer performance’ of the older age group manifested itself only in the popular pairs and therefore she argued that it was possible that participants by the age of 11 years had formed very definite ‘likes’ and ‘dislikes’ and the musical content in her test probably consisted of material that was more ‘disliked’ than ‘liked’ (p.25). A second strand to the research done by Castell was in the justification of responses. In the second part of her style sensitivity test she asked participants to try and write down what had made them decide whether the two pieces were from either the same or different pieces of music. Participant’s justifications were found to fall into one or more of a number of categories. One of the most common responses was to mention three specific aspects of the music, namely: a) Tempo / rhythm / speed; b) Instrumentation; c) Tone / pitch / melody. The interesting finding from this section of the study was that although participants’ had performed with greater accuracy in the popular pairs than in the classical ones, in their written responses, the participants appeared to have a more adequate vocabulary for describing classical music than for popular music. This finding led Castell to suggest that this was a reflection of the way in which the English music education system has a tendency to concentrate on classical music. It is an interesting anomaly in Castell’s research that participants more accurately identify ‘same’ and ‘different’ musical pairs in the popular excerpts and yet in terms of specific musical vocabulary, they appear to ‘know less’. This finding appears to suggest that perhaps participants are attending to more prominent cues (e.g. beat,

tempo) in the popular excerpts. These cues can be discerned and processed without the need to define, assimilate or describe in words what is being heard.

This primary attention to ‘beat’ was also noted by Flowers (1984). In her study, ‘tempo’ or ‘beat’ appeared to be the first category by which the younger participants judged the musical excerpts. This particular musical element would possibly be more prominent in the popular pairs than in the classical ones and therefore it could be expected that participants would perform more accurately on the popular pairs. It should be stressed here that the studies carried out by Flowers in relation to ‘tempo’ and ‘beat’ were carried out on younger children. It is therefore questionable as to whether or not these findings would be generalisable to older subjects. That is, whilst either prominent beat or tempo may explain some aspects of the test performance in the younger subjects, the poorer performance of the older subjects may be due to other issues, such as increased discrimination which may in fact lead to more uncertainty, or lack of clarity in the task. Therefore, this experiment attempts to further explore this problem by trying to systematically control the musical elements and further varying the stylistic divergence between adjacent musical periods.

9.1.2 Theoretical perspectives on musical style and participant preferences

Brittin and Sheldon (1995)

Brittin and Sheldon (1995) carried out a study on American students. Half the participants tested were music-majors whilst the other half were non-music majors. Participants were played examples of Western art music and their preferences were recorded either on a 10-point Likert scale or a Continuous Response Digital Interface (CRDI) which recorded their level of preference and how this changed continuously throughout the whole piece. The musical examples used keyboard, string orchestra and wind ensemble across Baroque, Romantic and 20th century sub-styles. The pace of musical samples was varied across the selection of samples, with half being slower than the other half. The main rationale for the study was to examine differences in preference ratings from non-music and music- majors in the two different response modes; that is to compare static responses (Likert scale) and continual responses (CRDI). They also explored listener responses to fast and slow examples across varied musical styles and different instrumentations. Results suggested some

difference between music and non-music majors. Music majors rated musical examples more highly overall than non-music majors but no difference was found between static and continuous response modes. In the non-music major results a significant difference existed between participants' responses in the continuous and static response modes. Continuous response participants averaged one Likert-scale rating higher than participants in the static response mode. Responses to changes in tempo were identical in both response modes and participants preferences for fast and slow examples were almost identical.

The difference between non-music majors' preference ratings in the continuous and static response modes was interesting. The authors suggested that non-music majors tended to react to the point of least liking:

“Perhaps less experienced listeners form their overall judgement judgements at the point of least liking, or perhaps they tend to remember their point of least liking”

(p.44)

Therefore in terms of static response modes, and style sensitivity tests as operationalised here may be considered a static response, there may be a significant difference in how and when preference judgements are made by musically experienced and non-experienced participants. Non-musical participants may react to their lowest point of liking in the specific musical example used in the test. Whereas more musically experienced participants may base their preference on a more global experience. This could extend to basing their preference judgement on their total experience of the whole style and not solely on the specific piece selected to represent that style in the test material.

Brittin and Sheldon also noted that all participants showed further interesting interactions between the instrumentation and the style of music, although this had not been a part of the original rationale for the study. Musical examples consisting solely of keyboard were least liked whereas string orchestra pieces were most preferred. The romantic period was the most popular style. A number of the earlier studies (Addessi et al. 1995; Hargreaves and North, 1999; and experiments one to four of the present study) had chosen to use 'piano only' musical examples for the classical pairs.

Whilst this may have indeed produced a more homogenous set of musical examples that 'varied only in style'; the findings of Brittin and Sheldon suggest that this may have correspondingly exacerbated the participants' tolerance of / preference for the classical examples. That is the less popular classical examples were additionally performed in a way that would be least preferred by participants.

Flowers (1983, 1984, 1990)

Flowers carried out a number of experiments that investigated attention paid to a wide range of musical elements by both elementary and college aged participants. Her first study (1983) was carried out on undergraduate participants and revealed that direct instruction in musical vocabulary appeared to affect participants' listening focus. Flowers (1984) attempted to build on the first study by extending the age range to include elementary aged children. In the first of two experiments, she compared third and fourth grade participants' descriptions of musical pieces with those descriptions made by college undergraduates. Participants were simply asked to listen to each piece of music and describe whatever they heard in the music. The results suggested that college students made a far greater number of responses across a wider area of musical elements than did younger participants. Flowers suggested that music education programs or general musical acculturation had some responsibility for younger participants' narrow focus on a small number of musical elements; that is, it was not simply a developmental feature. She suggested that younger participants were simply more familiar with frequently heard terms such as 'loud' and 'soft', whilst other musical terms such as 'high' and 'low' were less frequently used, at least in a musical context. Previous research (Hair, 1981) had suggested that younger participants could perform as accurately as college students in those areas of musical description with which they had some familiarity ('fast', 'slow', 'loud', 'soft'), and the participants in Flowers' study confirmed this. Younger participants mentioned dynamics and timbre (e.g. piano / violin) as accurately and frequently as did older participants. Flowers was therefore suggesting that the ability of younger children to perceive musical elements equalled that of much older participants. However, younger participants tended to only identify and describe those items to which either music educators typically drew most attention.

In the second part of the experiment, Flowers investigated the effect of teaching musical vocabulary and musical activities on children's written descriptions of music. In each experimental condition, participants experienced two music instruction periods in which a 'pair of musical terms' were taught in a wide variety of musical settings. The pairs of terms included were: 'piano' and 'forte'; 'staccato' and 'legato' and 'adagio' and 'allegro'. Following the periods of instruction participants were again asked to listen to the same musical excerpts and to describe the music. The findings of this investigation were interesting in that those participants who were taught the meaning of dynamics, articulation and tempo did appear to show an increased sensitivity to those elements as they occurred within the test pieces. Participants who were given tuition on timbre and extra-musical elements (i.e. images and analogies suggested by the music but external to it) did not show any corresponding increase in sensitivity to those particular elements. Flowers (1984) noted that participants showed an increased tendency to select and respond to the particular element upon which they had received instruction:

“ It would seem that choosing a manner of responding to 'what you hear' represented an either-or choice for most children; e.g. if they referred to articulation, then other types of description were generally omitted. Class vocabulary instruction had the effect of producing significantly more responses to specified musical elements; however, this gain was accompanied by some decrease in attention to other elements of music” (p.23)

In an additional study (Flowers, 1988) 62 elementary education majors described as having “little experience in the formal traditions of music listening” (p.22) were given music appreciation instruction along with a number of teaching demonstration lessons /presentations on musical appreciation. Each undergraduate was asked to locate their own group of young participants and instruct the group according to the presentations they had themselves received. The children were described as being “primary children aged from Pre-school to sixth grade” (p.23). The characteristics of the group of younger children, that is the size of the group, the musical background and the age of the participants was decided upon by each individual undergraduate.

The results of this third study led Flowers to conclude that older participants tended to describe music in a much fuller way than did young participants. As in the previous studies, younger participants still tended to use the exact words that had been used

within music instruction. In common with Castell's (1983) results it was found that tempo was the most frequently used dimension of musical discrimination and this was followed by dynamics. Flowers again remarked that younger participants attention to musical elements tended to be narrow and based on music vocabulary used in the instruction they had just received. She comments that if this narrow focus is truly what children hear; "... it is not surprising that they often remark, " that music sounds all the same to me" (p.32).

9.1.3 Analysis of musical style and participant preferences

Koroscik (1997) suggests that the level of understanding a person has in any specific domain, (in her case Painting), can be seen in terms of 'Novice' and 'Expert'.

" Experts possess a far more extensive knowledge base and much more effective knowledge-seeking strategies than novices"
(p.149)

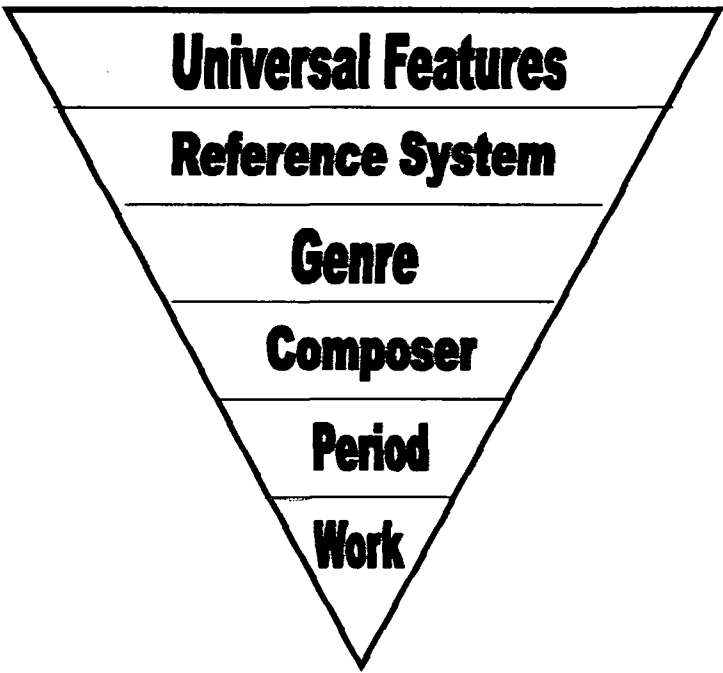
'Development' is defined as the changes which take place in the knowledge base, the knowledge seeking strategies and the learner's disposition in order to progress from the state of 'novice' to that of 'expert'.

Gardner (1972a) spoke of the development of style sensitivity as a progression from the 'untrained' (observer / listener) to the 'connoisseur'. In the untrained individual, a narrow and superficial number of elements are noticed but the 'connoisseur' may take a large number and range of cues into account. Gardner argued that " age of canvas, slant of signature, avoidance of certain motifs, jottings in a diary" (p.326) are all relevant to the true connoisseur when determining the style of a particular work of art. Similarly, participants in the present style sensitivity test may well draw upon a whole range of cues, such as production technique, instrumentation and volume depending on their knowledge base of the individual musical styles.

Nattiez.

Nattiez (1990) produced a hierarchy of stylistic phenomena that he presents in the form of an inverted pyramid (p.136). At the tip of the pyramid, Nattiez places the specific work under analysis. At this level the individual traits of the work, that is those traits that give the work its uniqueness from other works in the same group, are compared. At level two are those stylistic traits of the artist or composer, i.e. those

musical elements that fix the work within a specific period within the composer's life. Here Nattiez is referring, for example, to the musical traits that distinguish early from late Beethoven works. Level three relates to the musical traits which suggests the work of a particular composer as a whole. This would refer to those aspects of style which are specific to Beethoven (e.g. orchestration). Level four relates to stylistic elements within an entire musical genre. For example, the stylistic differences between two symphonies; one by Haydn and the other by Mahler (e.g. set keys and form of movements against varied key structure and varied movements). Level five consists of the tonal reference system in which a work can be placed; Purcell harmony compared to Schoenberg's serial tone system. Level six refers to the universal features of music, that is rhythm and variable pitches of sound.



Nattiez argues that stylistic analysis varies according to purpose. As an example, if the purpose of an analysis is to identify the composer of a piece; the level at which analysis will begin and the number of analytical levels required to fulfil this task will be very different from the analytical process required for a different purpose e.g. establishing whether or not a particular work is a forgery. Each analytical purpose dictates its own analytical approach; that is the entry level and the route taken by the analyst through other levels of analysis. The selected purpose for the analysis preconditions the level of stylistic relevance. (p. 135). That is, an analysis as to whether or not an unmade bed is a work of art would take place on the level of

‘Universal features’ and either remain there or it could move up through the layers. Whereas a question as to which period in the composers life the 7th. symphony of Beethoven belonged to, would start in either the ‘Period’ or ‘Work’ level and remain very much within those two levels. However, the full range of analysis suggested by Nattiez would only be available to an analyst who was either an ‘expert’ or ‘connoisseur’ in the specific subject domain under analysis. A ‘novice’ may undertake some analysis but the extent of their knowledge base would limit the number of levels on which the analysis could take place. As Koroscik points out:

“ The term, *novice* not only applies to most children, it also refers to any person, regardless of age, whose understandings are underdeveloped” (p.147)

Nattiez proposes that all analysis either implicitly or explicitly adopts one level or another, however the knowledge base of the individual making the analysis may be a greater influence as to the level or levels upon which the analysis is carried out

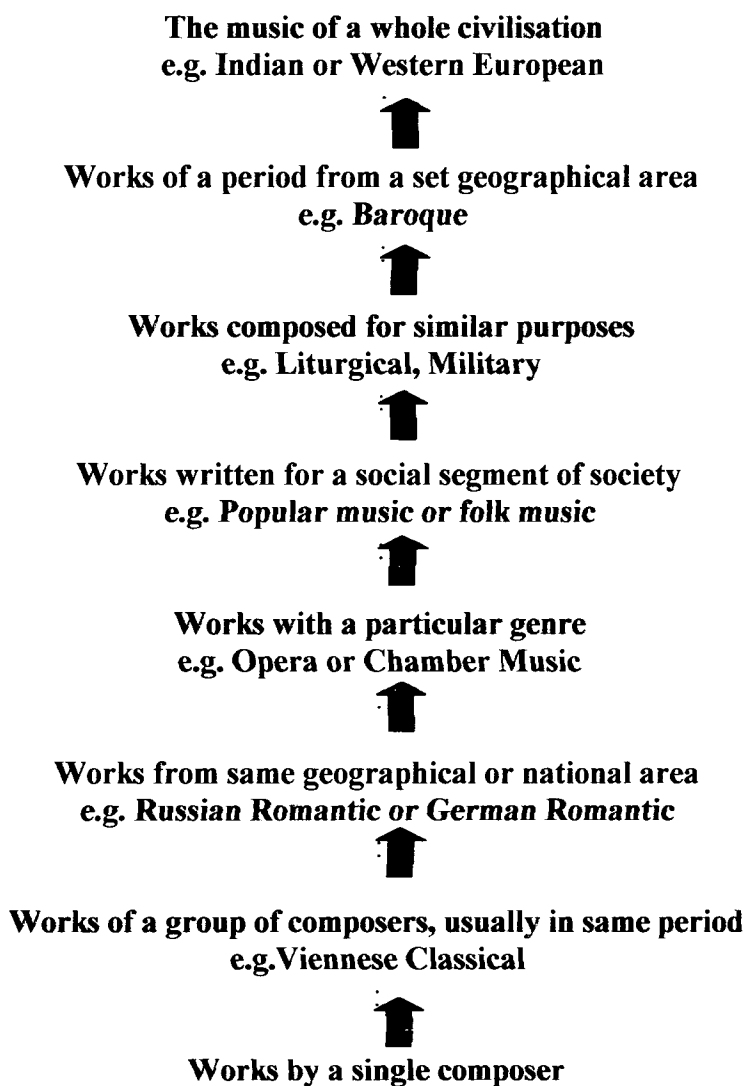
Meyer

Meyer (1989) discusses the issue of musical analysis and style and offers this definition:

“Style is a replication of patterning, whether in human behaviour or in the artefacts produced by human behaviour, which results from a series of choices made within some set of constraints” (p.3)

He argues that the ‘constraints’, [accepted rules, regulations, traditions] from which choices are made [key, pitch, contour, instrumentation and form] are learned and adopted as part of the historical / cultural circumstances of individuals and groups. Those who make choices [artists, composers, musicians] out of these ‘constraints’ are seldom the same people that make the rules. Constraints can be ignored to some degree, but for a style [the choices] to be ‘understood’ or for a style to have ‘meaning’, then a set of patternings that have shared meanings must be replicated in some way. That is; individuals who make ‘choices’ and replicate ‘patternings’ must remain to some extent within the constraints of the cultural system, or else shared meaning is not possible; at least not initially. Meyer claims that the goal of style analysis is to identify and formulate rules of musical patternings. Meyer proposes a similar model to that set out by Nattiez, however Meyer suggests a set of 8 levels

(p.39) of stylistic difference, compared to Nattiez’ five level model. Meyer’s model is therefore slightly more specific in some areas but does serve to emphasise far more the possible distance that may exist between popular and classical pair stylistic analysis. Meyers’ model is detailed here, however it has been inverted to place the individual work at the tip of the pyramid. This now matches Nattiez’ model:



As Meyer points out, and as Nattiez acknowledges, the potential number of divisions and sub-divisions is variable and will tend to be grouped according to the perspective of the investigator (analyst) or, to revert to Nattiez’ terminology, the analytical standpoint.

The work of Nattiez (1990) and Meyer (1989) provides an interesting idea as to how different cognitive processes can be operating on different musical pairs within the same test. When participants are asked to identify whether two pieces of music are

‘the same’ or ‘different’ they may not perform the task in the same way for all questions in the test. Simply because the same task request leads to a similar response does not indicate that the participant has used the same cognitive processes to move from the *task request* (Are these the same or different?) to the *task response* (Same or Different). In the style sensitivity task used here, participants may use very different cognitive processes according to the different musical excerpts – classical or popular. As Nattiez stated, each analytical task dictates its own level of analysis. In style sensitivity tests, as operationalised here, the level of analysis required to perform the discrimination between two classical pairs, may be very different from those required to discriminate between the popular pairs. For example, the difference between a Baroque piece and a classical piece may be complex and require analysis on Nattiez’ levels one, two and three; that is Work, Composer and Period. Whilst the same request for analysis on a piece of Metal may require only one level; namely that of ‘Genre’.

In this musical experiment some effort has been made to manipulate the musical material used in the test so that both classical and popular pairs require the same level of Nattiez’ stylistic analysis. The experiment simply seeks to discover whether or not some form of systematic manipulation of musical elements can affect the level on which participants perceive, discriminate and analyse the musical samples.

9.1.4 Summary: The present study

In summary, previous research suggests that there are a number of interesting questions to be addressed. Firstly, what are the influences that operate within the style sensitivity test that may account for the consistent pattern of different scores obtained between the popular and classical pairs. Do participants score more accurately on the popular pairs because they are familiar with and ‘recognise the patternings’ of those musical styles, in other words they arrive at the answer through ‘stylistic competence’; that is, a competency in that style. Alternatively, are there simply a greater number of prominent musical cues in the popular extracts that can be attended to aurally; that is subjects are sensitive to differences in the musical cues without any internalised knowledge of the musical ‘patternings’ of that style.

There are certainly a number of possible reasons why subjects appear to perform with less sensitivity in the classical pairs than on the popular pairs. It is possible:

-That the main effect for age reported by Castell (1983); that is the decline in open earedness, could be extended to musical style. Castell noted that subjects' mean scores decreased with age; younger subjects (10/11 years old) performed with greater apparent accuracy than older subjects (13/14 years old). It was also noted that the apparent decrease in style sensitivity occurred within the popular examples. Castell had argued therefore, that subjects' tolerance towards popular musical styles decreased with age as subjects had developed more clearly defined musical preferences in line with peer group pressure. However it is also possible to expect that many of the classical examples would also fall beyond subjects' immediate preferences. Therefore, any apparent decrease in performance due to a decline in 'open-earedness' could be expected to equally affect both classical and popular genres.

-That "social and cultural factors exert a powerful influence on any potential cognitive determinants of style sensitivity" (Hargreaves and North, 1999, p.200).

Participant's responses are influenced by a number of extraneous factors that cannot be controlled within the test.

-That in two of the studies (Addessi et al., 1995; Hargreaves and North, 1999 and Experiments one to four of this thesis), some reaction occurs through the 'piano only' examples used as test material. (Brittin and Sheldon, 1995).

-That participants' music education programmes reinforce a limited number of musical elements through which participants group and analyse musical samples (Flowers, 1983, 1984, 1988, 1990).

-Judgements of preference and therefore tolerance may be made differently depending on participants' musical experience. (Brittin and Sheldon, 1990)

The main question to be asked here therefore is whether any one, or any combination of the above factors fully explains the frequent performance differences between

participants in all conditions and across all age groups, on the popular and classical pairs?

The second group of questions concerns the cognitive process which participants utilise in order to process the musical material involved in the style sensitivity test.

- Do participants score better on the popular pairs because they are familiar with and recognise the ‘patternings’ of that musical style, in other words real ‘stylistic competence’

or

- Do participants score consistently higher on the popular pairs because there are a greater number of prominent cues they can attend to, in other words a simple matching task?

9.2 Selection of musical examples

Condition Co (control)

Within this control condition, the musical excerpts in the contrasting pairs were identical to those already used in the previous experiments one to four. The samples consisted of the four classical styles: Baroque, Classical, Romantic and Modern; and the four popular styles: Indie, Grunge, Blues and Thrash. The same tape recording and reproduction equipment was used in this condition as had been used in experiments one to four of this present study.

Condition Cc (Contrast)

In this condition, all procedures and methods were identical to those in experiments one to four although the level of contrast was changed. The way in which this was done is described below. Within the popular styles, one example was taken from each of the following: Grunge, Ska, Metal and Jazz and the classical examples were taken from Baroque, Classical, Romantic and Modern periods. Each individual style was contrasted against itself and against each of the other three styles in the same category (i.e. A-A, A-B, A-C, A-D, B-B, B-C etc.). This gave ten combinations: namely Baroque with Baroque; Baroque with Classical; Baroque with Romantic; Baroque with Modern; Classical with Classical; Classical with Romantic; Classical with Modern;

Romantic with Romantic; Romantic with Modern and Modern with Modern. The Popular music examples were paired in the same way giving a further ten examples: Grunge with Grunge; Grunge with Ska; Grunge with Heavy Metal; Grunge with Jazz; Ska with Ska; Ska with Heavy Metal; Ska with Jazz; Heavy Metal with Heavy Metal ; Heavy Metal with Jazz and Jazz with Jazz.

Condition Ce (eras)

In this condition, the style sensitivity test was identical in method and procedure to the other two conditions. The difference in this condition was that the chronological time scale between musical periods in the classical section was increased. In this condition the previous four classical styles were replaced with musical examples from Renaissance, late Classical, late Romantic and 20th. century modern. The popular pairs, method of recording and procedures were kept identical to the other two conditions. This condition had the effect of varying the level of stylistic divergence between the classical pairs by using pairs that were not from adjacent musical eras.

9.3 Measurement of stylistic divergence

Tafuri et al. (1994) outlined the difficulty of isolating and controlling the musical traits within the musical material used in the style sensitivity test. In an attempt explore this issue and for the purpose of establishing a greater level of stylistic divergence between popular and classical pairs in the musical material used in Condition Cc, a simple and broad - grained method of measurement was developed in an attempt to establish an informal quantitative process for assessing the approximate level of 'contrast' between any two musical samples.

First, a number of musical components that could be identified as prominent differences between both musical pairs, and therefore possibly be audiated by participants and used in their decision making process were identified. A number of researchers have explored how individuals compare two short pieces of melodic-rhythmic pattern with respect to "the range of potential distortions" possible between the two excerpts. (Fiske, 1992. p.363). For example, where two contrasting pairs varied in key, Fiske found that ease of discrimination between the two samples varied as a function of the distance between the two keys. Subjects found it easier to discriminate between two keys that

were distantly related; ease of discrimination varied as a function of key distance (Dowling and Fujitani, 1971; Dowling, 1978).

Initially, the 'musical categories' identified by subjects in Castell (1983) and Flowers' (1984) research, were adopted. Ultimately, eight categories of musical component were identified as possible sources of 'potential distortion'. Each of these categories are defined more fully below. See also Longuet-Higgins, (1976), Simon and Sumner, (1968) and Fiske, (1992) for a full review.

1 & 2. Tempo and Rhythm.

The tempo and the rhythm of the piece are here treated as two separate components. Within the literature, there is some debate surrounding exactly what aspects of rhythm / tempo / metre is heard by individuals. Different individuals either hear rhythm and tempo differently or they group rhythmic sounds together and represent them in different ways. (see for example Radocy and Boyle, 1997; Bamberger, 1979,)

1. Tempo.

Two musical examples may differ or contrast by means of the tempo. That is simply the pace at which the music progresses regardless of the rhythm or the metre. Any rhythm or metre may be played at different tempo (speed). In some modern pieces, the tempo may be difficult to identify easily within the short test example or may it change within a short period of time. (Fraisie, 1982)

2. Rhythm.

Two contrasting pieces could vary in time signature (e.g. 6/8 against a 4/4) and yet the prominent metre (the beat) may be identical in both pieces and played at the same pace. That is, a 6/8 rhythm can be counted in 2 (beat); similarly a 4/4 time signature can also be counted in 2. Again, in some modern pieces, the rhythm may be difficult to identify easily within the short example used or it may change within a short period of listening. Yet within most popular pieces the rhythm and meter tend to vary little, are pronounced and easy to detect. (see 'figural' and 'metric' Bamberger, 1975; Deutsch, 1982; Fraisse, 1982)

3. Key.

Two contrasting pieces could be presented in a different key. No examples were used in the test in which major and minor keys were contrasted. Examples simply used different major or minor keys. For the measurement procedure used here, a decision was made that when two keys varied by less than one whole tone (eg.F major & F# major), then these two keys would be treated as the same. (Dowling and Fujitani, 1971; Dowling, 1978; Shepard, 1982)

4. Instrumentation.

The main instrument carrying any main theme; the number of instruments playing; the combination of instruments playing and the tone of a similar instrument could all produce an audible contrast. In the case of the same instrument being prominent, attention was also given to the tone production of each instrument for example as in a Baroque piece and a Romantic piece with the former played on a period instrument. (Risset & Wessel, 1982)

5. Instrumental Production.

The instrumental production of the two pieces could give audible clues by providing a contrast. In the popular examples, this could be present between two electric guitars using different 'effects' boxes, or the level of sound to noise ratio inherent within the production itself. In classical examples this could be a recording made in a studio and a recording made in a church or concert hall.

6. Style of Playing.

Contrast could be present between the style in which a common instrument was being played. For example, a prominent violin played in Baroque style using strict tempo, accurate bowing and little or no vibrato presents a stark contrast with the same instrument played in a Romantic style with rubato playing, lush bowing and full vibrato. Whilst most children may be unable to describe such differences, it is possible that they may notice 'something' different between the styles.

7. Recording Production

The difference in production methods of the recording may present another contrast. For example a strong difference in the overall sound is present between a baroque piece recorded in a studio and the same piece recorded in an auditorium or church.

8. Volume

There may be a large or small contrast in the volume of the playing between the two examples. This may be due to a marked crescendo rather than an increase in the number of instruments. Participants may simply notice that one excerpt 'sounds louder or quieter' than the other. For the purpose of this experiment, the volume between two contrasting pairs was not assessed by use of signal to noise ratio on the recording equipment nor was it based on the output meter of the reproduction equipment. A simple but sensitive sound to voltage was purpose built. The output volume was registered on a 12 point LED bar display.

It is accepted that this approach recognises a specific definition of 'style' and how it may contrast. This approach is course grained but effective and appropriate to the research question. That is, the purpose of the experiment is to explore if manipulation of the musical material used in the test through varying contrast level and increasing the chronological time between eras, can affect participant test performance. It does not seek to explore further how or which musical traits may cause greatest influence. It is argued firstly that any one, or a combination of the musical prominent musical cues outlined above could be present within the test samples used and may in turn be presenting audible cues upon which the participants were basing their decisions. Secondly, it is argued that whilst some definitions of 'style' may well encompass a greater range of phenomena than the above noted components, there is no proof that children would be able to detect this more complex product.

9.4 The assessment of musical samples: Condition Cc (contrast)

The components noted and defined above were each given the equal value of 1. The same classical eras were used, namely; Baroque, Classical, Romantic and Modern but a change was made to the popular examples. As in the previous study, Metal and Grunge were kept the same but Indie was replaced by Ska, simply because by its very

nature, Indie cannot be defined as a set style of music but as a movement of independent musicians.

The precise musical excerpts used in this condition were chosen in a different way from those used in the control group. In common with the control group examples, each example was again 15 seconds in duration and there was a 15 second period of silence in between the two pieces. No vocal pieces were used. No example contained the start or finish of a piece of music and where the two examples came from the same piece, the second example did not continue directly on from completion point of the first piece. All musical material was recorded on identical equipment to those samples used in the control group and an identical quality and make of magnetic tape was used. The order in which the contrasting pairs appeared on the tape was identical to the order used in the control group tape and the introduction and response sheet were identical.

The selection and the recording of musical samples were carried out in the following way. The classical samples were selected first and a quantitative value was assigned to each contrasting pair. This quantitative value was determined by a value of 1 being given where each of the musical components stated above was found to be different between the two contrasting pieces. For example, where the classical and baroque contrasting pair differed in instrumentation, key, style, volume and tempo; the pair was given a contrast value of 5. This procedure was carried out on all the classical combinations and each classical pair was given a contrast value. A full list of the ten contrasting pairs was created ranking the various pairs in ascending order according to their contrast value. Where two or more pairs had an identical value, pairs were listed in alphabetical order; designated by the first style in the pair. This was not considered to have any methodological implications as this list was not involved in the placing the various pairs in test order, but simply a means of ensuring that all pair values were matched equally with a popular contrasting pair.

The popular contrasting pairs were then determined through a similar process. The first popular excerpt was recorded onto tape. The second excerpt was then selected and recorded and a contrast level was determined. Reference was then made to the classical 'list' and a classical pair of equal contrast value was removed from the list.

Obviously, as recording continued, more values were removed. The last three examples of the popular pairs had to be ‘customised’; that is they had to fulfil a specific contrast value. In these instances the first excerpt was recorded and a prolonged search took place until a second excerpt was found which would produce the required contrast value. The second excerpt was therefore selected not only according to the required style but an example of that style which created a specific level of contrast. The overall pattern of contrast values was therefore matched between the classical and the popular pairs.

The order in which participants experienced the classical and popular pairs was the same across all conditions.

9.5 The assessment of musical material: Condition Ce (eras)

It has already been noted by Gardner (1973a) that within the classical contrasting pairs, participants’ optimum scores were achieved between styles from non-adjacent periods. For example, there would be greater chance of participant success between a contrasting pair containing a Baroque example and a Modern example, (distant eras) and less chance of participant success between a contrasting pair containing a Baroque and a Classical example (adjacent eras). It is argued here that participants in the Gardner study were possibly responding to the phenomenon of greater contrast between pairs. That is, the contrast level between the adjacent eras was less pronounced than between the non-adjacent eras. As was reported earlier, Addessi et al. (1995) produced a contrasting result. It was therefore decided to explore participant performance in style sensitivity where the contrast level between the Classical contrasting pairs was increased by moving the musical styles further apart chronologically into non-adjacent categories. The four classical music styles were therefore changed from Baroque, Classical, Romantic and Modern to Renaissance, Mid-classical, Late Romantic and Late modern. The hypothesis was that by selecting classical styles from non-adjacent periods would lead to participants achieving equivalent scores in the classical and the popular musical pairs. A full list of the contrasting classical pieces used is given in Appendix 2. The popular pairs used in the test were identical to those used in condition two. All musical examples were recorded on identical tape and equipment and all other procedures and methods were

identical to those used in experiments one to four. Similarly, the order in which participants experienced the various contrasting pairs was identical to the other two conditions.

9.6 Design and procedure

In summary, the experiment incorporated three conditions C (control), Cc (contrast level) and Ce (eras). In condition C, participants received the same style sensitivity test and procedure already described in experiment one. The test material, introduction and method were used according to the procedures already set out in chapter 4.

In condition Cc, the level of contrast between the musical pairs was assessed. Classical musical excerpts were selected which gave a similar contrast value to the popular pairs, as assessed by the method stated earlier. In condition Ce, the classical musical samples were selected from non-adjacent musical eras as a further method of increasing the level of contrast between the pieces within the contrasting pairs.

The order of the classical and popular pairs remained the same across all three conditions and following the pattern already established in previous experiments, participants' papers were initially checked through and any papers in which the participant had clearly not understood the task were withdrawn. Following this initial removal of papers and in accordance with procedures carried out in experiments one to four of this study, the class teacher withdrew an appropriate number of responses at random in order to make the number of responses equal in all conditions.

9.7 Schools

The experiment was carried out in a number of different schools and employed a total population of 398 participants: 7/8 year-olds, 104 participants; 10/11 year-olds, 104 participants; 13/14 year-olds, 94 participants and 15/16 year-olds, 94 participants. In the two elementary age groups, the participants were taken from one large primary school in North Surrey. The number and variety of musical activities that were offered by the school and the amount of music and musical activities that the children experienced within the average school week were judged to be similar to those schools used in previous experiments. The school offered some choral work, a small

amount of violin and woodwind tuition, and some recorder work. In year 8 and year 10, the 13/14 years old participants and the 15/16 years old participants came from a comprehensive school in North Surrey with a 5-form entry. All participants followed the national curriculum music syllabus and around 19% of participants received instrumental tuition. Participants in both year groups and in all three conditions took the test in small mixed class groups during an afternoon period.

9.8 Participants

7/8 years old age group.

In this group, 104 children came from one school with three parallel mixed ability classes in each year group. Each full class group was split into three groups of approximately ten children. As in the previous experiment, participants were initially placed into groups according to friendship with some subsequent minor adjustments made in order to balance male and female participants and to maintain mixed ability groups. Each group was designated by the name of one pupil in the group. The groups attended the test in an order determined by the alphabetical order of the group name. All groups were tested in the school library with the particular condition selected by one participant selecting a marked token. Participants in each condition comprised three groups of approximately ten participants from each of the three class groups. Final numbers of participants for each condition were as follows; Condition C: 33 participants; Condition Cc: 36 participants and Condition Ce: 35 participants. Responses were further reduced to 90, that is 30 per condition by random withdrawal by the class teacher in the manner already described.

10/11 years old age group.

In this age group participants were taken from the same primary schools as the 7/8-year-old participants. Testing again took place in the school library on a different afternoon to the younger participants. The groups for each condition were constituted by the method already described for the younger participants. Each condition comprised of three groups of approximately ten children taken from the three classes in the year group. All groups were judged by the teachers concerned to be mixed ability. The choice of experimental condition each group received was decided in the manner already described. The total number of participants tested in the 10/11 year-

old age group was 104. There were 35 participants in condition C; 35 participants in condition Cc and 34 participants in condition Ce.

Responses were further reduced to 90, that is 30 responses per condition by random withdrawal by the class teacher in the manner already described.

13/14-year-old and 15/16 year-old age groups.

The 13/14 year-old age group comprised 94 participants. In each of the three conditions participants experienced the test in whole class groups. Conditions C had 31 participants, condition Cc had 31 participants and condition Ce had 32 participants. Following the random removal of papers by class teachers a total number of 90 responses were used, 30 responses from each condition. The choice of experimental condition was made by the class teacher who selected one of the three tapes. In the 15/16 year-old age groups a total of 96 participants experienced the three conditions in whole class groups. Condition C had 32 participants; condition Cc had 32 participants and condition Ce had 32 participants. Selection of participants and choice of condition was achieved through an identical method to that employed in the 13/14 year-old age group.

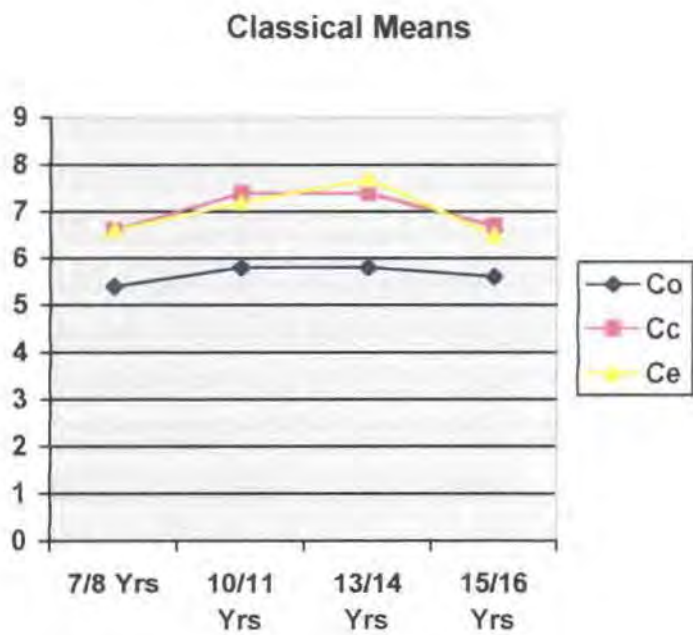
Following the process already established previously in experiments one to four, response sheets were marked with a score of 1 being given were the participant correctly identified the two excerpts as being taken from the same or a different piece of music. Where the participant made an incorrect response, no score was given. Each participant had two final scores; one score, up to a maximum of 10 for the classical pairs and the same for the popular pairs

9.9 Results

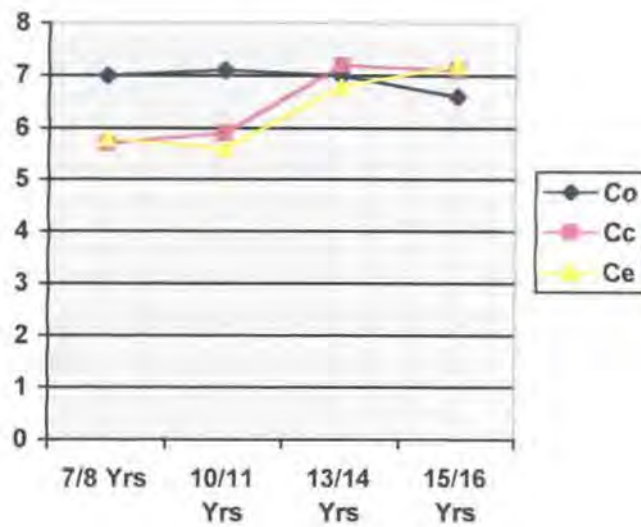
A three way mixed analyses of variance was carried out on the mean scores with one within-participants factor (musical style: classical / popular extracts) and two between subjects factors - age group [4 levels] x condition. As all subjects experienced both the classical and the popular pairs, this was a repeated measures analysis. Further analysis of both the individual classical and popular score means was carried out separately by the use of post hoc Tukey HSD tests.

The main effect for age was statistically significant ($F [3,348] = 10.97, p = 0.00$); likewise, the main effect for experimental condition was also significant. ($F [2,348] = 7.87, p = 0.00$) whereas the main effect for style was not ($F [1,348] = 0.22, p = 0.64$). For the three main two - way interactions, there was found to be significant interaction between experimental condition and age ($F [6,348] = 2.19, p = 0.04$); and also between musical style and condition ($F [2,348] = 39.75, p = 0.00$). Likewise the interaction between age and style was also found to be significant ($F [3,348] = 7.00, p = 0.00$). Lastly, the interaction between age, condition and style was found to be significant ($F [6,348] = 3.97, p = 0.001$)

The interaction means are plotted in Fig. 9.1



Popular Means



9. 10 Discussion

General Points:

The main hypothesis was that participants completing the style sensitivity test in the two conditions with increased level of contrast (Cc and Ce), should perform more accurately than those participants who completed the standard test material in the control group. This hypothesis was found to be correct with participants in conditions Cc and Ce obtaining significantly higher means for the classical pairs than for the popular pairs. Secondly, in common with previous research, the level of accuracy achieved in the style sensitivity test by even the youngest participants was high with even the youngest participants achieving total scores well above those expected through chance.

Classical pairs

As in the previous experiment and in common with the findings of Gardner and Castell, the main effect for age was statistically significant. Participants mean scores increased with age; the 13/14 year-old participants performing most accurately. Style sensitivity then appeared to decrease with the eldest participants achieving lower mean scores than the 13/14 year-old participants. Subsequent analysis of means using Tukey HSD tests showed the differences to be significant between the 10/11-year-old

age group and the 13/14 year-old participants who achieved mean scores of 6.79 and 6.98 respectively ($p= 0.01$). The difference in scores between the 13/14 year-old participants and the 15/16-year-old participants was also significant with mean scores of 6.98 and 6.29 respectively ($p= 0.01$).

The main effect for experimental condition was also statistically significant. Participants in Condition C performed at higher levels in the popular pairs than the classical ones; that is style sensitivity appeared to increase amongst the excerpts that were most familiar. Participants in both Cc and Ce conditions demonstrated a contrasting result, obtaining higher means in the classical pairs than those participants in the control condition. This result suggests that manipulation of various musical elements within the excerpts used in the test material may affect subject performance in style sensitivity tests which are operationalised in this way.

The main effect of musical style was not found to be significant and this confirms the hypothesis that increased levels of stylistic divergence can remove the apparent difference between participants' performance on familiar and unfamiliar genres. This result suggests that apparent significant difference between popular and classical means, as obtained in previous research and experiments, may result from varied musical elements within the test material and not necessarily from social influences.

Popular pairs

As in the classical pairs, the main effect for age reached statistical significance for the popular pairs. Scores in all three conditions also tended to follow the pattern achieved in most of the previous experiments with style sensitivity appearing to improve dramatically between the 7/8-year-old participants and the 13/14-year-old participants with the eldest group achieving lower scores, but not to any significant level. Subsequent analysis of means using a Tukey HSD tests showed the difference in scores between the 7/8-year-old and 13/14-year-old participants to be significant ($p=0.000$), as was the difference between the 7/8-year-old, 10/11 year-old participants and the 15/16-year-old participant ($p=0.00$ and $p=0.00$) respectively.

Taken as a whole the results again suggest that age related changes do occur in participants performance in style sensitivity tests, as operationalised here. Style sensitivity does appear to increase between the ages of 7/8 and 13/14 years old and this is followed by an apparent decrease in performance in the 15/16-year-old participants. This age related pattern of a significant increase, followed by an apparent decrease was maintained in this experiment in spite of the changes occurring in other patterns. Similarly, the interaction between age and musical style again suggests that different developmental trends may exist between familiar and unfamiliar genres. That is, in spite of the overall mean scores improving and in spite of any difference between the classical and popular means being removed, the overall developmental trends remain different between the two styles of music.

As reported earlier, Gardner (1973a) had found that his participants were more sensitive to those styles from distant musical eras. Accordingly, the findings here do appear to add support to his suggestion that the level of contrast was greater between these more distant musical periods. However, the findings do not accord with those of Addessi et al (1995) who had argued that style sensitivity did not vary as a function of distance in time, but as a function of distance from those 'musical patternings' the participants had internalised. In the particular instance of their participants, they suggested that Mozart and Baroque music had been internalised far more. Therefore, the findings here may not be totally contrary to the results obtained by Addessi. et al. That is, any number of different experiences between the two populations could produce differences in which styles the participants had internalised the most. Music education programmes, changes in musical styles favoured by the media, musical fashions favoured by prominent films, can all influence the dominant classical diet that participants receive. Therefore, it may be the case that in both studies, style sensitivity did vary as a function of distance from those styles the participants had more internalised. However it may be that those styles were in fact, very different.

The results also add support to the comment made by Tafuri et al. (1994); that it is difficult to both isolate and to control the different and numerous musical traits within the musical excerpts used in a style sensitivity test as operationalised here. As the results of this experiment suggests; these elements may outweigh other influences. It

was suggested earlier that participants in some age groups might be reacting to or identifying prominent cues within the music and not necessarily identifying musical patternings. However, controlling and isolating those prominent elements is a complex and difficult issue. As Gardner (1973a) pointed out when he used a voice as a prominent figure against a musical ground, many younger participants failed to notice the vocal line but based their decision on other musical elements.

Similarly, Flowers (1984) pointed out that younger participants not only identified and grouped musical material together in a very different fashion to adults or older participants, but grouped musical elements together in different ways. For example, participants described tempo changes (fast and slow) as ‘happy’ and ‘sad’. Therefore, one tune in a minor key and one tune in a major key but both played at the same tempo, were classed as ‘the same’. Therefore adults who attempt to isolate and control variables, as was attempted here, may find that the different variables they identify, control and isolate are not necessarily those that younger participants appreciate.

9.11 Conclusions

The main experimental hypothesis was that by manipulating the level of stylistic divergence between the classical and the popular pairs (as measured according to the procedure outlined above); that the significant difference previously obtained between participants performance in the classical and popular pairs would no longer be present. The contrast level was varied in two ways and the hypothesis was proven.

The results appeared to support the notion that artefactual elements within the musical examples used, substantially influence the scores that participants achieve in both the classical and popular pairs. The ramifications of this result suggest that further style sensitivity tests using a similar format may need to establish some greater control over musical elements within the samples used. Certainly the result suggests that a number of previously obtained results may in fact have been, at least to some extent, a result of musical features within the samples selected. However, one developmental trend appears to remain. Regardless of the actual scores achieved within the classical and popular pairs, whether they are similar or significantly different, style sensitivity does

appear to improve within the younger age groups and there is a corresponding decrease in apparent sensitivity around the 15/16 year old age group.

9.12 Summary

The experiments reported so far in this thesis have attempted to explore further interesting interactions which may impact on the perceived participant performance within the style sensitivity test, as operationalised here. Secondly, they have explored further the reliability and validity of the instrument of measurement which has been used in a similar fashion in all previous studies carried out into the development of style sensitivity in children. The previous studies in question can certainly be subjected to a number of criticisms. Firstly, a very small number of studies have been carried out in this area with each of the studies involving a relatively small research population. Secondly, no previous study has attempted to explore the validity of the test instrument. All previous studies have operationalised style sensitivity in an almost identical way with no further attempt made to expand either the definition of style sensitivity or the age range, neither has any previous study attempted to explore the large variety of test variables which can affect participants' performance. Yet in spite of this, a number of substantial claims have been made based upon the collected data.

The five experiments carried out and reported so far, have suggested that participants performance on the style sensitivity test, at least as operationalised here, can vary as a function of who presents the test, the educational programme they are / have been involved in, the test introduction and the actual segments of the musical pieces used in the test itself. So although Hargreaves and North (1997) may in fact be correct when they suggest that, "It is probably impossible to disentangle the cognitive aspects of responses to musical styles from their social and affective components" (p.200), it is probable that the list of possible influences to disentangle probably goes well beyond social and affective components.

Castell (1982) suggested that the poorer performance of the older subjects in the poplar pairs, was due the younger children being more 'open – eared' or tolerant to the musical samples used, yet she made no attempt to explain the more significant

difference between the participants' performance in the classical and popular pairs, which should, at least in theory, have been more sensitive to the 'open-earedness' effect. That is, participants were least tolerant of the classical styles. Tafuri et al. (1994) suggested that the poorer performance of some of the older children could be attributed to their recent change from primary to secondary education and that children internalised the patternings of classical music first due to their prominent use in the media.

Essentially, what the first five studies in the present study suggest is that the instrument of measurement may well lack both validity and reliability, with participants' performance being greatly influenced by a whole range of characteristics of the test environment. Therefore, it is certainly a possibility to say (after Rosenthal, 1976) that the personal characteristics of the presenter, the test material, the test environment, the educational experience and the test instructions did in some way determine the participants' response and we must certainly therefore, hold our knowledge more lightly for that.

Similarly, it is probably impossible to disentangle not only the cognitive, social and affective components but also the interactions between what has been learned formally and informally and what has been taught and that which has been understood.

Chapter Ten

Experiment six

Assessing musical style sensitivity in pre-schoolers: an exploratory study

“Language can never adequately render the cosmic symbolism of music, because music stands in symbolic relation to the primordial contradiction and primordial pain in the heart of the primal unity, and therefore symbolizes a sphere which is beyond and prior to all phenomena”
(Nietzsche, 1967, p.55)

This chapter reports on the rationale for, and the development of a new procedure for testing musical style sensitivity in pre-schoolers and presents some of the previous literature which influenced the design of the new test procedure. In the first section, some background to the study is presented and this is followed by a brief review of some previous literature on musical capacities in early infancy. There have been a number of previous studies which have highlighted methodological issues surrounding the testing of musical abilities in young children. These methodological issues both informed and influenced the design of the new procedure and a brief review of these studies is given. The next section describes the development of the new testing procedure for style sensitivity in pre-schoolers and finally, an exploratory study on the effectiveness of the new procedure was carried out on a number of participants and the results of this study are written up and discussed.

10.1 Background

Previous studies on the development of style sensitivity, including the first five experiments in this thesis, without exception have been carried out on children within a similar age range namely between 6 and 16 years old, (18 and 21 in Castell's (1983) study). One reason for the concentration of studies in this age range has inevitably been the unsuitability of the testing procedure for younger children and therefore no previous study has ever employed any participants below the age of 6 years old. Previous style sensitivity tests have required participants to respond with varying amounts of written work, and whilst this has had the advantage of allowing a large amount of data to be collected quickly and efficiently, the procedure has also had the disadvantage of being relatively long and intensive, frequently taking between 20 and

30 minutes to administer and complete. Whilst older participants may not find this task an arduous one, it is probably an inappropriate way to explore style sensitivity in children below six years of age.

Gardner (1973a) commented that even the youngest participants (age 6) in his study appeared to possess high levels of sensitivity to musical style, achieving scores well above those expected through chance. If this finding is correct, it is certainly valid to explore style sensitivity in younger age groups and to question what developmental trends exist in children between the ages of 0 and 6 years old. Just as social and educational factors may influence preferences and tolerances of older children, further studies would need to explore what are the cultural, educational and musical experiences that influence the early development of style sensitivity.

“ Musicologists may be interested to note that present-day
infancy research pays increasing attention to the early
perception, processing, and production of melodic sounds”
(Papousek, 1996. p.37)

If some level of style sensitivity is present in participants aged 6 years, then research on the development of style sensitivity in the pre-six year old child could be a most fruitful area of study. For this research to take place, a number of more age appropriate testing procedure would be required.

One possible criticism of previous studies into the development of style sensitivity is that they have all incorporated a similar testing procedure. Although the two studies by Tafuri et al. (1994) and Hargreaves and North (1999) used and compared both triadic and pairwise tasks, all previous studies have incorporated a similar lengthy procedure with two (or three) musical excerpts separated by a period of silence, being played to participants who are required to respond with varying amounts of written work. Any inherent weakness or bias in that testing procedure from any of the sources reported in chapter three of this thesis, could therefore be reflected throughout all the results reported by each of the previous studies.

In chapter one of this thesis, some discussion took place of the advantages and disadvantages of using either short or prolonged musical excerpts as aesthetic stimulus in studies of listening to music. For example, writers such as Aiello and Sloboda (1994) have commented that issues relating to long and short-term memory,

and not just musicality, may influence results in tests which use prolonged musical excerpts. Whilst the excerpts used in the first five studies of this thesis, and in previous studies may not be termed 'prolonged', at least in the sense used by Aiello, Sloboda and others, neither can they be termed short extracts in that they are not simply 5 to 8 note sequences. A substantial amount of musical information is present in many of the classical excerpts and the period of silence between the two excerpts may create a memory issue for the participant. Previous studies have typically used musical excerpts of an average of 15 seconds in length followed by a period of silence. This period of silence between excerpts has varied. Castell (1983) used 5 seconds, experiments one to five of this study used 15 seconds, Gardner (1973a) used 30 seconds. The remaining two studies did not record the duration of the interval between the excerpts, so the excerpts may even have been presented consecutively with little or no space separating them.

Salzer (1952) commented that the understanding of a musical piece may depend on the listener connecting similar musical segments. In the case of Gardner's (1973a) study, participants would not only have to identify those segments but also retain them for 30 or more seconds; that is six times longer than participants in Castell's study. Participants' failure to perform accurately on some of the musical pairs therefore, may stem from a difficulty in identifying and memorising musical cues and not from a lack of sensitivity to musical styles. It is therefore valid to question the importance and the role which participant memory may play in the procedure used by previous studies.

Cortazzi (1993), although writing about the design of research questionnaires has raised an issue which could be equally pertinent to style sensitivity tests, at least as operationalised in previous studies. Cortazzi expressed the view that a research 'questionnaire' should be viewed as a complete linguistic event; the questionnaire is a compound event with each question affecting subsequent questions. It is impossible to present an individual participant with a question and then completely erase the experience or impact of that question. Therefore, Cortazzi contends that the experience of each individual question can be the sum of :

- a) The individual's reaction to the question. This could be complex and contain such elements as an emotional response (e.g. anger at the subject matter / negative feelings), or a desire to please the author of the questionnaire.

and

- b) The actual response of the individual to the question. This may be a desire to hide a truth or represent an image the individual wishes to portray to the author.

The sum of this complex experience can impact on all subsequent questions throughout the questionnaire with the result that the individual's reaction to the very last question on the questionnaire, can be a compound reaction to all previous questions. Any one, or combination of initial questions may impact on all subsequent questions. Applying this idea to style sensitivity tests, a participant experiencing the first pair of musical excerpts in a style sensitivity test which features a 30 second period of silence between the two excerpts, may well experience some difficulty in memorising the musical cues. This may impact on all subsequent excerpts with the listener ignoring the musical cues they would normally attend to, and attending instead to other cues which they may feel are easier to remember.

Participants expecting a longer period of silence between excerpts may well concentrate on far less musical material than participants expecting a short period between excerpts. The period of silence may direct the amount of, or the type of musical cue the listener concentrates on.

The period of silence between musical excerpts which has been incorporated into previous style sensitivity tests, may have exerted some influence on the accuracy with which participants performed. In addition to issues surrounding memory span, younger participants may also be subject to distraction. Even relatively short periods of 10 seconds of silence between musical excerpts can be sufficient to allow for distraction or a break in concentration in much younger children and this in turn would affect their performance.

Therefore any new testing procedure would benefit from addressing the issues of memory span, concentration and possible distraction from the task. This experiment

goes some way towards that by allowing participants some element of choice over the time span between musical excerpts and also by providing a more novel and concrete task.

10.2 Research on musicality in infancy

A number of areas of research suggest that infants develop musical perception skills throughout childhood regardless of whether or not they receive any formal musical training. Chang and Trehub (1977) discovered that infants of six months of age could discern differences in auditory patterns such as variations in the time period between notes resulting in a different rhythm. Thorpe, Trehub, Morrongiello and Bull (1988) found that infants spontaneously grouped sound sequences together, forming temporal structures that are similar to those sequences formed by adults. They suggested that this might be an indication that even within the first year of life, young children are able to hear and discriminate temporal properties in the musical sequences they hear and impose this organising structure on other less familiar sound sequences. Trehub and Thorpe (1989) suggested that infants are able to recognise the similarity in rhythm across sound sequences in spite of ongoing variation in frequencies and tempo or the rate of presentation.

Chang and Trehub (1977) further suggested that infants were able to discern differences in tonal patterns. Infants, having learned a series of tones, were able to identify when the same tones were played in a different sequence. Further studies by Morrongiello (1986) found some support for these findings suggesting that infants were able to discern melodic pitches and remember their temporal order. Infants were also found to respond differentially to changes in the order of tones that affected overall melodic contour and to those changes that did not. This suggested that infants are both able to discern differences and also similarities between different tonal patterns. This tendency to focus on the melodic contour of an unknown piece rather than other features of the melody has been observed both in pre-schoolers and adults. (Bartlett and Dowling, 1980; Dowling, 1978). Similarly, Trehub, Thorpe and Morrongiello (1987) found that 9 – 11 month old infants could identify melodic sequences of similar contour consisting of different frequencies and different melodic intervals.

Some of these studies have also discovered more extended developmental processes in which constant accommodation of new musical material is required. Trehub, Cohen, Thorpe and Morrongiello, (1986) discovered that infants performed at a lower level than younger pre school participants on musical discrimination task. Participants were played two melodies, one 'correct' and one 'incorrect'. The incorrect melody contained a semi tone pitch change to one note. That is, the second melody would be considered to contain a 'wrong note' according to the harmonic rules governing western classical music.

Pre-schoolers appeared to be far better at discriminating the incorrect melody than the older participants, who responded equally to both correct and incorrect melodies. The authors suggested this was a result of the greater exposure to 'correct' melodies than the older participants, who presumably may have had more exposure to a wider range of musical styles.

Papousek (1996a) has reviewed the biological and cultural origins of early musicality. In his review of the literature, Papousek argues that predispositions for competences in musical behaviour develop from both genetic and environmental factors with many of them appearing before birth. He argues that in past research, infant musicality has been too often linked with emotionality because music can mediate human feeling. However, drawing on research into listening abilities in both the unborn infant and early infancy, he suggests a number of psychological competences exist that are common to both language (communication) development and musicality; both are the result of some form of evolutionary selection. His thesis is that musical development in infancy should be explored more closely with the corresponding development in language.

The competences he outlines suggest that high levels of auditory discrimination are present even prior to birth. Papousek contends that a predisposition for musical discrimination, along with the ability to combine words and playful modifications to acoustic elements, have been selected by evolution in a particularly advantageous combination.

“ Musical elements participate in the process of communicative development very early; in fact, recent research has suggested that

they pave the way to linguistic capacities earlier than phonetic elements can do” (p.43)

During the first two to three months, interactions between adult care givers and infants tend to be characterised by rich melodic experiences (Papousek and Papousek, 1981) and musical qualities. Constant repetitions and social reinforcements give the young infant the ability to anticipate, control and answer the elementary musical components. The capacities of the newborn infant for auditory discrimination tasks are great. New born infants appear to be able to recognise a familiar voice, story or melody (De Caspar and Fifer, 1980; DeCaspar and Spence, 1986; Cooper and Aslin, 1989); near term foetuses can distinguish between male and female voices, (Granier-Deferre et al., 1992) and two day old newborns already show a pre-disposition for their own native language, (Moon et al., 1993). All these tasks occur through inter-uterine experience and require high levels of auditory discrimination.

Therefore, as Papousek (1996a) points out, “ the human newborn enters the social world with an almost adult-like hearing organ -the inner ear (p.43). Infants are far more sensitive than was previously thought (Olsho et al., 1987) and the discrimination of musical elements may play a vital role in early communication and early language development. The list of musical abilities and awareness of musicality in infants’ lives and capabilities is growing fast and the use of new and novel measures of infant abilities are constantly increasing (Deliege and Sloboda, 1996).

As has been noted, Gardner (1973a) commented on the high ability of even the youngest participants (aged six years) in his style sensitivity test. These results suggested that style sensitivity developed well before the age of his youngest participants. In view of the research reported briefly above, this is not surprising.

10.3 Methodological issues

Kellett (2000) has pointed out that listening to music requires sustained concentration, a skill that many young children find difficult. She also suggests that the standard research / assessment strategies often used on this age group do not allow a full appreciation of the level of musical ability in very young children. That is, the true musical ability level of young children is often underestimated due to either a lack of concentration skills, concentration span and also the intervention and requirement of

language skills in test procedures. Kellett further notes that musical research often uses the judgements of professional adult musicians to norm reference the child's response rather than exploring the child response for its own intrinsic value. Although her work was not actually carried out on preschoolers, the points raised by Kellett are certainly pertinent.

Kellett (2000) carried out her study on six to eight year old children. Each child was given a set of colours, patterns and textures. Having listened to an extract of music, each child was required to tick a response box according to the colour, the pattern and the texture they best thought suited the music. The experiment was carried out over a prolonged period of time (one full term) and was designed to build self-esteem amongst participants and to encourage them to make more responses to musical extracts. Kellett commented that the practice of recording responses non-verbally had a number of significant effects. Firstly, children who were usually more reticent in committing answers to paper were far more willing to respond by way of a non-verbal task. Secondly, participants who were usually dominated in other tasks by more verbally able peers were found to take a full part in the activity.

The use of verbal test instructions, no matter how simple, can also be a difficult issue. Ward (1984) wished to explore the concept of 'consonance' and 'dissonance' in young children. He claimed that presenting seven-year-old children with three two-note musical chords and asking participants to make a judgement as to which chord they felt was the odd one out helped participants display their judgements and concepts of 'dissonance' and 'consonance' without having to verbalise their reasoning. This approach certainly relieves the participants of having to make a verbal response but still requires the participant to understand the terminology, 'odd one out', at least in the same sense as the experimenter. Flowers (1984) commented that younger participants appeared to judge any form of written activity as being a very different task from the identical aural / oral activity carried out only moments earlier. That is, where children were taught certain musical terms for the purposes of a test requiring a written response, their understanding of those terms was not transferred to a verbal response test. Terwogt and Van Grinsven (1991) observed that in musical tests requiring language activity, very young children often attempted to resolve their

frustration with their own inadequate language skills by, 'singing, whistling or moving' i.e. by expressing their ideas through an alternative form of activity.

Kastner and Crowder (1990) attempted to link participants' ideas of the emotional content of music with appropriate linguistic labels. In their study they linked the language elements of the test with schematic facial expressions. For example words such as 'happy', 'sad' and 'frightened' were accompanied by a drawing of a face representing that particular expression. Wood and Coltman (1998), although referring to mathematical research in young children noted that often the youngest participants spent so long having to verbalise or remember the full task required of them that they forgot what they needed to say.

Greer, Dorow and Hanser (1973) studied participants in grades 2 and 3. Their research explored the effect that teaching discrimination skills had on the musical preferences of young children. The experiment utilised both classical and rock music examples. Participants receiving the music discrimination training were taught to match four pictures of musical instruments to musical selections featuring different instrumental timbres. The results suggested that participants' musical preferences did change as a result of tuition. Participants receiving the musical discrimination tuition changed their musical preferences from rock music to classical music. The authors do not suggest why this change should occur but there are a number of possible explanations. Participants may have changed their preferences in line with those of the teacher through some form of 'prestige effect'. Alternatively, participants may have found the classical music more interesting because they were able to complete the discrimination task and identify the instrument.

Schuckert and McDonald (1968) suggested several reasons why pre-school music should be taken very seriously. They suggested that music education at this level enhanced the capacity for self-expression (Heffernan and Todd, 1960); promoted the ability to listen intelligently (Garretson, 1966); facilitated social behaviour (Newman, 1966; Taubman, 1958) and introduced children to cultural traditions (Hill, 1966). They further noted that despite the apparent importance of pre-school musical education in forming later musical habits and attitudes, very few investigations have been carried out in this age group.

With specific reference to style sensitivity, Hargreaves (1999) in recounting an impromptu song improvisation session between himself and his two sons (aged 4 and 5) notes that even at this age, both boys were able to mimic stereotypical and recognisable 12-bar blues conventions that:

“include ‘blues notes’ such as minor thirds and sevenths: and the phrasing and general delivery of the genre, including blues ‘breaks’ at appropriate points in between verses” (p.31)

He concludes that “ it is quite striking that this particular genre or ‘cultural frame’ is so clearly established in a pair of British pre-schoolers”.

A review of the research suggests that children much younger than 6 years old, having received appropriate exposure to the musical system of their society, not only have considerable musical abilities but also the ability to demonstrate these under appropriate and suitable conditions. It therefore seems not just legitimate and valid to question how and at what age young participants can be said to become sensitive to musical style, but vital. Similarly, the role played in this process by formal and informal music education is an equally important area for new research.

Therefore an alternative procedure was designed and tested. A full description of the method is presented later, but briefly, in the new procedure participants were presented with 2 sets of musical excerpts; 4 classical and 4 popular. Four minutes of each individual excerpt was recorded onto two short identical tape cassettes; 2 minutes of one musical excerpt on one tape and the following 2 minutes on a second tape. The cassette tapes were then mixed up and given to the participants who were required to listen to, and match all the cassette tapes into pairs. The two cassette tapes in the matched pair contained what the child thought was the same piece of music. The main differences in the new procedure are that:

- participants are tested individually, not as part of a group.
- the required response is a concrete ‘matching’ activity and not a written response.
- participants may vary the time of silence between hearing excerpts.
- participants hear different sections of each excerpt. That is, unlike the previous experiments, participants may hear up to 4 minutes of music from

each excerpt as opposed to the short extract they heard in the previous experiments. They also have control over the time period between the samples they hear and can hear the samples as often or for as long as they wish.

10.4 Design and procedure

Two main prerequisites were set down for the new procedure. Firstly, that it should be equally appropriate for a wide age span with minimal adaptation: maintaining the interest and concentration levels of very young participants, but also be sufficiently novel to be of interest to older participants. Secondly, in view of the studies reported briefly in the previous section, a non-verbal response mode was felt to be most appropriate.

10.5 Musical examples

For the classical pairs, the same musical styles and excerpts were used as were used in experiments one to four of this present study; namely Baroque, Classical, Romantic and Modern. However, the popular excerpts required a substantial change. In order to comply with the terms agreed between the writer and the schools, no musical excerpt was to contain any vocal lyric. This was done in order to avoid children hearing anything which could be construed as either unsuitable or inappropriate lyrics and cause offence. Previous excerpts had utilised the short instrumental ‘break’ within a song. The new procedure required four popular pieces representing four different, individual styles and also containing over four minutes of continuous instrumental music. This was challenging in that the styles previously represented rarely incorporated extended instrumental music. Therefore four contrasting excerpts were selected from four different artists; namely Vanessa Mae, David Arkenstone, Ultravox and Yanni. The four excerpts chosen were distinctive and all used electronic instruments. Whilst each artist could not be said to be a representative of a particular style, it could be said that the individual excerpts used were being performed in a particular distinctive style. Excerpts were chosen to closely resemble the four styles used in previous experiments namely; electric Jazz, Metal / Thrash, Grunge and Ska. Full details of the final selection of excerpts appear in Appendix Three.

The experiment utilised 8 individual pieces of music: 4 classical and 4 popular pieces with each piece representing a distinct musical style. A total of 4 minutes from each individual piece was recorded onto two cassette tapes; that is 2 minutes on one

cassette tape and the subsequent 2 minutes onto the second cassette tape. Sixteen new, plain black cassettes of tape were used. The cassettes were made especially for the experiment and contained only 2 minutes of recording time. Each cassette looked identical with no distinguishing features. No recorded excerpt started at the beginning of the piece, nor did it include the ending of any piece. Each piece was recorded in such a way that two minutes of the piece was recorded first onto one tape. Then the music was stopped and the next two minutes recorded onto the second tape. Following the recording, all eight classical cassettes were marked with a yellow dot in one corner of the casing. All eight popular cassettes were each marked with a red dot.

A pre-test matching activity was also incorporated and this consisted of a set of 8 white 5x5 cms. cards. A rough outline drawing of a fish, a person, a bus and a tree were drawn onto the cards. However, the drawing was done so that each figure required two of the cards.

10.6 Participants

A group of 15 three and four year old participants were selected to pilot the new testing procedure. All participants attended a nursery unit attached to a village primary school in Surrey. The unit offered a range of musical activities including music and movement, percussion instruments and singing occurred on a daily basis. Participants were tested in the school office or quiet area of the nursery. A familiar adult was always close by. Participants were brought to the test area and introduced to the tester as follows:

“ Mr. Marshall, this is (child’s name).

(Child’s name) this is Mr. Marshall and you are going to help him to listen to some music. O.K.?

I will come and collect you when you have finished”

To allow the participants to settle in, they were asked a number of simple questions, namely;

-Do you like listening to music?

-What is your favourite music?

-Do you hear a lot of music at home?

Participants were shown the tape player and asked if they knew what it was.

Some participants showed a desire to talk and report at length on their own familiarity with tapes and recorders and music; however some simply nodded or just answered 'yes'

" So, you know you can use this to play tapes and listen to music. So here is a tape; and there are two players; this one is your side and this one is my side. The tape goes in this way... ..and this button makes it play... like that". (Demonstrated)

" Now, can you put the tape in for me and switch it on so that it plays?".

Participants were then given the opportunity to become acquainted with this process. No time limit was set on this and the exercise continued until participants demonstrated an ability to operate the tape player. Almost all children, even the quiet ones, showed a keen desire to show their ability at this.

"OK, now you are very good at that. Is the music too loud for you? ... (Volume is checked). Now we can leave that for a moment. Let's have a look at these cards.

[One half of one figure was placed on the table].

Now, I drew this man (fish, tree, bus), but there was not enough space to draw it all, so I drew some more on another card. This one. So, if I put them together, they are a pair and they make the picture. See, this picture is different from this picture, but they are from the same drawing... I just did not have enough room. OK.

Now here is half of a... (train, fish, person, bus)

[Another card was picked up]

Can you find the other half?

Good, now can we find the other half to these two pictures.

[This was continued until child had matched all four cards]

Now if I drop these cards on the floor, they all get mixed up and we have to match them up again. Do you understand that?

Well, I have these tapes. And I put some music on one tape, but there was too much music, so I had to put the other half on a second tape. Just like on the drawing. So, I had half the music on one tape, and half the music on the other tape.

So let me tell you about something that I did today. I picked up all these tapes and dropped them all over the floor!

So now, I don't know which tapes go together in a pair... like on the pictures. So, we can choose one tape and then I want you to see if you can find the tape with the other half of the music. Let's have a try.

In this box are all my yellow tapes, and in this box are all my red tapes and I dropped them all. So, shall we do the yellow or the red tapes first?

[Child selects colour of tape and tapes are tipped out onto desk].

So, we are going to match them up again.

Would you like to pick one of these tapes... any one, and give it to me”

Now, I will put this tape in my side and we can listen to some it. [Tape plays for approx. 8-10 seconds.]

OK. Now you pick up a tape, any one... and play it in your side and see if it matches mine or not.

Now you can listen to my tape as often as you would like to, so if you forget or want to check, then we can just listen to it again.

At this point, the child placed a tape in the machine and played a section of the music.

They continued to play each of the other tapes until they agreed on a match.

When the child suggested they had found the matching excerpt, both excerpts were played one immediately following the other. The child was asked:

‘Do you think they are the same?’”

This was done in order to address the issue of memory which, it was suggested earlier may have influenced the results of previous studies. If children wished to check back with other excerpts, they were allowed to do so. If the child still claimed the excerpts matched; both cassette tapes were placed together at one side of the table.

The child was then told,

“OK, good. Well done, we have now matched two of the tapes, shall we do another one now. Pick another tape for me, and you have to find the matching tape”

This process was repeated until all the excerpts were matched. The process was then repeated with the second batch of tapes containing the excerpts from the second genre.

10.7 Assessment procedure

A log was made of the order in which each participant listened to the two musical samples. Each style was denoted by the initial letter of the style name; e.g. B- Baroque, R- Romantic, J-Jazz. For example, the log for a participant trying to find part two of a baroque excerpt would be as follows:

Ruth: Aged 4 years

Excerpt	Time			Action Taken by Participant
	A	B	C	
B		x		Baroque excerpt selected and played

C	x	Classical excerpt played but not compared with original baroque excerpt
R.	x	Romantic excerpt played but not compared with original baroque excerpt
M	x	Modern excerpt played but not compared with original baroque excerpt
B – B original	x	Baroque excerpt played and compared with
B – C	x	Baroque compared with Classical again.
B – B (C)	x	Baroque excerpt played and compared with original and the decision made and was correct.

The recorded log of events produced several pieces of information. It denoted whether participants made an instant and confident decision or whether they continued to change their minds. It also denoted the number of times a participant had to hear the original piece and those pairs they wished to hear compared several times. The log detailed above for Ruth, records that she made quick decisions once she had heard all four excerpts. It records that she felt there was no need to hear the modern and romantic pieces a second time, and that she was confident with her decision having checked it one further time against the classical excerpt.

All final decisions were treated in an identical fashion, that is with a verbal response of encouragement such as ‘ Good’ or ‘Well done’.

A broad measure was also taken of the time each participant decided to spend listening to each of the musical sample being played. This measurement was not precise but consisted of marking one of three time categories, namely:

- Approximately 0 – 2 seconds. (A)
- Approximately 2 – 6 seconds. (B)
- Anything over 6 seconds. (C)

The log for Ruth shows that she did not feel she required hearing large quantities of the excerpts, and the final decision on the match took place very quickly, in less than two seconds.

10.8 Results and discussion

A score of 1 was given for each correct comparison made and all final results were expressed as a percentage. As each individual made a different number of comparisons, it was more accurate to represent the total correct comparisons as a percentage of the total number of comparisons made. Of the 15 participants tested in this age group 5 identified between 85% and 95% correctly; 2 participants 69%; 3 participants 48% and 5 participants between 0 – 9%. With reference to the time taken on each comparison; 81% of decisions were made within the 0-2 second time category and 19% of decisions were made in the 2-6 second time category. No participants took longer than 6 seconds to arrive at a decision. Lastly, a one way ANOVA was carried out on the style sensitivity scores. The main effect for musical style (classical and popular) was found to be significant ($F [7,14] = 23.63, p=0.001$), with participants performing more accurately on the popular pairs than on the classical pairs.

It is notable that a high percentage of these preschool participants were able to make reasoned judgements and to discriminate between the musical excerpts, although whether or not they were truly discriminating between 'styles' is questionable. Participants tended to be quiet and although all were tested in the presence of another more familiar adult, several children nodded rather than responding verbally. One very young girl did not wish to operate the tape recorder and during the second batch of tapes, she simply nodded to every piece of music as it was played and left hurriedly at the end. Perhaps further research on this age group would benefit from a procedure more in line with that described by Kellett (2000) and described earlier. A more accurate measure of children's style discrimination abilities could be obtainable if a series of tests were to be carried out over a prolonged period of time, possibly as part of a musical activity period by a more familiar test presenter or class teacher within a familiar environment. This would help to clarify whether or not several of the participants could genuinely not perform accurately on the activity or whether they simply wanted to remove themselves from the test situation as quickly as possible.

That having been said, almost all participants very much enjoyed the task, as evidenced by their willingness and insistence in operating the tape themselves, but

even the participants who could perform the task well began to show signs of tiredness towards the end. A number of typical transcripts are shown below:

Tommy: Age 4. Very confident and very animated.

1. B
M
R
B- (1 second correct match); 'That's it; it goes ne ne ne ne ne'
During this explanation, Tommy stretched his eyebrows and blinked his eyes in time to the 'ne ne' ne' ne'.
2. M
R
M- (1 second. correct match). 'Yes, now this one'.
3. C
C - Ha! 'do do do do do do' (He sings and pretends to play a violin although no violin is prominent in the music)
4. R
R - Yesssssss; that one goes bm bm bm bm bm bm

All responses were made very quickly and with confidence.

Ellen: Aged 4

1. M
C
B
R
Pauses: {Prompt- "Shall we try another one; you choose"}
C
B
M- "I think that one".
2. C- "Hmmm.....it's faster"
R "oh-oh; not that"
B- "Must be thissssssssss....one"
C- "That one - can I hear that one?"
C-C "Yeh, that one"
3. B "yep"
R "yep"
4. R-B "that one"

Parsons (1987) commented that in describing developmental stages, many writers tended to concentrate on what younger age groups could not do, as opposed to describing what they could do (p.26). Gardner (1973a) offers support for this view. He pointed out that his youngest participants showed no awareness of specific dimensions in the music and when questioned about the reasons behind their choices, they appeared to be unable to give any reasons for their decisions. He suggested that, "the lack of specificity and precision reflected limited vocabulary rather than global perception" (p.73). Gardner goes on to suggest that an important consideration in the perception of styles is whether or not two styles 'feel' different and he further argues that many participants perhaps create metaphors in order to share their experience of the works, and these metaphors may not necessarily be verbal. In keeping with previous studies, participants appeared to experience some difficulty in giving verbal reasons for their decisions; in fact some hardly spoke at all. Yet many of the participants appeared to demonstrate some "feel" for the musical excerpts in movements and facial expressions. In the example given above, Tommy mimed playing a violin along with the classical music although no prominent violin appeared on the musical excerpt. Yet he somehow associated something about the musical style with playing the violin. Often participants would add some movement to the music. This may be a small movement such as clenching and unclenching a fist or nodding their head. One child giggled and squeezed her shoulders to both jazz music excerpts, which she correctly matched. Yet the same child seemed unwilling or unable to offer any verbal explanation for her decision.

These responses appear to add some support for the previous research noted earlier. Kellett (2000) noted a number of benefits from her methodology which are equally applicable here. She noted that the use of an individual non verbal / written test procedure encouraged more reticent children to take part and make their responses. Certainly in this exploratory study, most participants were willing to take part and make a response. However, even those participants who were unwilling to say very much did take part and many completed the task well and displayed a high level of accuracy. Similarly, Terwogt and Van Grinsven (1991) had commented on the way in which younger participants resolved their verbal frustrations by expressing their 'ideas' through alternative forms of activity such as movement, singing or whistling. This phenomenon was partially supported through observation. However, again many

participants were much quieter and far less animated within the test situation than in their own classroom group. Again, more longitudinal studies in which participants became far more familiar with the test presenter could reveal many more interesting ways in which younger participants did express their perceptions of musical excerpts through 'alternative forms of activity'.

10.9 Conclusions

This study has explored the development and the use of a new form of testing procedure which may be applicable to pre-schoolers. Participants were given the opportunity to hear each musical extract as often as they wished; to repeat any combination; to vary the time between excerpts and to respond to the task in a practical way. The results suggested that some sensitivity to musical style is present in children as young as four years old with up to 50 percent of the participants in this pilot study displaying an ability to discriminate and match some prominent musical cues from a good number of musical styles.

The results also suggest that preschoolers found the new form of testing procedure to be both manageable and enjoyable and but there are a number of ways in which it may be criticised.

One major criticism would be that although the procedure was successful in providing a mode of response which did not require any form of written work, it is still heavily dependent upon language. Certainly, amongst younger participants it is hard to establish whether or not those participants who performed poorly, did so because they did not understand the task, or because they were not able to discriminate between the musical excerpts. However, as a good number of the participants did perform well on the task, it does appear that many children in this age group are capable of understanding the idea behind the task. That having been said, certainly future studies could benefit from a number of further changes. Following the example set by Kellett (2000), linguistic content of the test could be reduced by including the style sensitivity test as part of a whole series of multi media activities or a small project involving matching a broad range of 'same' and 'different' artefacts (shapes, colours, paintings, drawings, movements etc.), this would reinforce the concept of the task in a number of ways and possibly reduce the amount of language content required.

Similarly, younger participants would benefit from a tester who was known to them. Future studies could include more longitudinal studies involving regular contact with the tester would create a more relaxing test situation. It was outlined earlier how younger participants appeared to be more susceptible to examiner influence, especially in a one to one test situation. Certainly, future studies would need to provide adequate control for this.

This exploratory study has involved a new form of test procedure which appeared to provide some interesting results on a small population of pre-schoolers. Initial results suggest that young children are capable of discriminating between musical cues and styles and performing well in understanding and carrying out a task of this type. The results also suggest that young participants are a fruitful, interesting and greatly under researched population in terms of the development of style sensitivity. Further research should accommodate more creative procedures and response modes which will allow children in this age group to fully display their capabilities. Certainly pre-school children appear to be able to give a broad range of aesthetic responses to different musical stimuli and any future study would benefit from noting and recording all their reactions. As Gardner (1973a) stated:

“ It is perhaps most reasonable not to penalize subjects who are unable to articulate their reactions and to attribute certain subjects’ abilities to describe the effects of music to their greater *imaginativeness rather than to heightened musical understanding*”
(p.75)

Chapter Eleven

Summary and conclusions

11.1 Introduction

This research has explored the development of musical style sensitivity in children aged 4 to 16 years and there were a number of reasons for choosing to research this particular area of music psychology. Firstly, relative to the amount of research carried out into other areas of musical activity, studies involving the process of listening to music have been under-represented. More specifically, studies exploring the development of style sensitivity have been equally rare since the interesting and promising start made by Gardner in 1973a. Secondly, more recent studies, both generally within the social sciences and specifically within the psychology of music, have begun to highlight the importance of giving greater consideration to the social context in which musical events take place. Lastly, discriminating between musical styles is a psychological process which is entered into many times each day, either consciously or sub-consciously, by the majority of humans and the development of this ability is a valid and interesting area for research and this thesis has attempted to extend what is known about this particular musical activity.

The present research has attempted to review, augment and build on the results of the four previous studies. The experiments reported in this thesis have further explored the developmental trends in both classical and popular musical genres by systematically varying the social context in which the tests were taken; the musical test material included in the style sensitivity tests and by extending the age range of participants to include children as young as 4 years old.

In this final section, I shall firstly present a brief summary of the results obtained by each of the six experiments carried out and reported in this thesis (11.2). In the subsequent section (11.3), I shall propose a developmental model and argue that style sensitivity and stylistic competence are not interchangeable labels describing the same mental function, but can be viewed as two distinct processes. This model draws

together a number of theoretical ideas and attempts to describe participants' decision making process in tests of style sensitivity, at least as operationalised in this thesis.

Lastly, the final section (11.4) presents a brief summary and a number of implications for further research.

11.2 Summary of findings

The first two experiments investigated the influences and effects of a number of social variables present within the testing environment. The first experiment explored the effects of changing the test environment. Previous research (Killian, 1990; Kuhn, 1975) had suggested that children might behave differently in music lessons and so the experiment attempted to study further, whether or not participants who experienced the style sensitivity test within the context of a music lesson would perform more accurately than children who took the test in a non-musical environment. It was hypothesised that children in the music lesson may become more tolerant of, and 'open-eared' towards all musical styles through one or a combination of influences, namely; prototypicality of the music, the natural reinforcing effect of music and musical background, and the typicality of testing environment.

The results of the experiment suggested that the musical environment did not appear to influence participants performance in style sensitivity test, at least as operationalised here. It was acknowledged that both tests were carried out within the powerful environment of the school. Therefore it is accepted that the influence of school may have affected both conditions. In retrospect, comparison with a further condition beyond the school environment could have created some further interesting and contrasting results. However, the main effect for age was found to be a significant with different developmental trends occurring for the classical and popular pairs. In common with previous studies, the main effect for musical style was found to be significant with participants appearing to be more sensitive when judging musical styles which were most familiar to them. Overall, the results of experiment one suggested that whilst the social context in which musical examples are heard might indeed affect some of the listener's aesthetic responses to that music; it did not appear to affect their performance accuracy in style sensitivity tests, at least as operationalised here.

Experiment two manipulated a number of attributes of the test presenter, namely; age, and cultural allegiance as exhibited by dress code, hairstyle and language code. The hypothesis for the experiment was that participants within the two conditions involving younger test presenters would achieve higher mean scores than participants in the control group. Previous research had suggested that race (kinship), gender, the age of the experimenter and the perceived prestige of the test presenter could affect participants' test performance. The results of the experiment suggested that subject age and musical style did have considerable effect on a participants' performance. The main effect for experimental condition was not statistically significant, although further examination of participants' mean scores did suggest some interaction between participants' age and gender and the gender of the presenter. This interaction appears to support the findings of Lutey and Copeland (1982) who also suggested that:

"age and not just sex of the children influenced whether they did better with a male or female tester". (p.132)

The results of experiment two suggested that 'who' presents the actual music may be as important an influence on subject performance as 'what' music is actually presented. The results of this experiment suggest that participants may be influenced by different examiner attributes at different ages and this could be a fruitful and interesting area for further research. Previous research had also suggested that experimenter attributes might vary as a function of test material. That is, not all test contents are equally susceptible to examiner attributes and therefore not all types of test content are sufficiently sensitive measures of examiner effect. However, the results of this experiment did suggest that tests of style sensitivity, at least as operationalised here, are an adequate and sufficiently sensitive instrument of measurement of the influence of examiner attributes. As discussed earlier, there are probably several reasons why the characteristics of the presenter should influence the test performance of the individual. Similarly, no one single effect can be said to account for all the variations in test performance. It is probable that a variety of emotions were present in the interactions between the participants and the test presenter. Examples of these would be attraction, jealousy, identification and rejection of current teenage cultural values; and it is likely that some emotions would be more pronounced in some age groups rather than in others.

Experiment three contrasted three groups of participants, one group in the United Kingdom and two groups in different regions of the United States. Participants' musical experiences varied in two main ways. Firstly, participants in each of the two countries experienced a different form of musical education program with the United States, in some ways, according music education a higher status than the United Kingdom. Secondly, there was some difference in the commercial 'musical diet' which participants in each of the three communities experienced. The differences resulted from two main sources. Firstly, there was a difference in the number and prominence of musical styles featured regularly in local media stations and secondly, there were differences in the relative status accorded to certain types of music by adults and by the general geographical community (e.g. Country and Western music was more prominent and acceptable in Iowa).

The results of the experiment suggested that whilst the American music education program, with its greater status and opportunities, may affect certain musical skills; no statistically significant differences in sensitivity to musical styles were apparent. The results appeared to offer some support for the argument put forward by Geake (1999) and Morrongiello (1992), that style sensitivity may not be one single musical skill, but the process of being sensitive to musical styles may depend on a combination of very different skills. Some of these skills may be unaffected by formal music education.

The results further suggested that although the difference in the apparent levels of style sensitivity between the three groups was not statistically significant, some differences in the developmental trends for classical and popular genres did exist. Participants from the U.K. and New Hampshire experienced similar exposure to popular culture and these groups showed similar developmental trends in the popular genres. Participants from Iowa and New Hampshire experienced similar formal music education programmes and showed similar developmental trends in the classical genres. These differences might suggest that the development of sensitivity to different musical genres is dependent on different musical experiences and influences. This experiment suggests that the development of sensitivity to classical styles might be more dependent on formal music education whilst the development of sensitivity to popular styles might be more dependent on cultural experiences. More detailed

studies would be required to address the more precise role played by, and the interactions between formal teaching and informal learning. Certainly, broadening the research population to include further differences between the schools and their geographical locations would be valuable here. The findings do however suggest that the instrument of measurement was sufficiently sensitive to identify minor differences in the populations which may have been due to teaching and learning.

Experiment four explored the effects of motivation. Participants in one condition were rewarded whilst the second condition was competitive. This experiment produced interesting and contrary results to those expected. It was hypothesised that participants in the reward and competitive conditions would perform more accurately than those participants in the control condition. However, subjects in these two conditions showed a tendency to perform with less accuracy than those in the control group. It was argued that participants had taken a longer period of time to make their decision as to whether the two musical excerpts did in fact come from the same or a different piece of music. In the case of younger participants, it was proposed that their decisions were no longer based on an immediate discrimination between prominent musical elements in the two pieces, but that they had attempted to listen more carefully and therefore relied upon musical knowledge of styles that they did not possess. That is, the increase in desire to do well changed the decision making process from the simple task of being sensitive to prominent musical elements, into more complicated cognitive process which, as yet, they were incapable of doing well.

In the older participants it was argued that the competition simply added to the already greater caution and uncertainty that was typical of this age group in other studies. That is, they had an even greater tendency to be over cautious and judge styles as being different. Further studies would need to establish more precisely the role played by the offer of a reward, the increase in motivation or whether or not the actual variation in test instructions had influenced the participants' performance.

Experiment five attempted to explore the effects of varying the musical material by incorporating two ways of increasing stylistic divergence. This was achieved by attempting to control the contrasting elements in both the classical and popular pairs,

and by increasing the chronological time scale between the classical musical periods. The results showed that increasing the stylistic divergence between the musical styles did appear to remove the usual difference between the classical and the popular scores. Participants tended to score either equally in both classical and popular pairs or score higher on the classical pairs. These results contrasted sharply with those achieved in previous studies.

Lastly, in experiment six, an exploratory study was carried out with the aim of developing a different form of test methodology which would enable much younger children to demonstrate their level of sensitivity to musical styles. The study drew on the findings of previous studies on the musical abilities of young children and overall, the results suggested that participants as young as three years old were quite capable of making accurate discriminating judgements about musical styles. Although the study was exploratory, the results suggested that not only were young children capable of making discriminating judgements, but also that the new test procedure was sufficiently sensitive to measure these discriminations. Similarly, future studies could both extend and validate the new procedure by combining and contrasting results from both the old and the new test format.

11.3 Implications for further research

In retrospect, there are a number of aspects of this research which could have been treated differently and this in turn, presents a number of implications for future research. The more general points are dealt with initially and this is followed by a number of more specific points.

In looking critically at the research design, it is now apparent that the research populations were very much under utilised. Many of the schools and teaching staff expressed great interest in the project. Many wanted to know ‘how the children had done’ and several asked for a copy of any publications to come out of the study. Many schools felt that they were not achieving their potential in music and valued the opportunity to take part in the study. One Headteacher when thanked for her help simply stated: “ Well, I view it as being as much what the school can get out of it and

not just what we have done for you". There is no doubt that this interest could have been harnessed and used to produce very different data on style sensitivity and certainly an opportunity for at least two or more longer term studies exploring the relationship between educational programmes and the development of style sensitivity has been missed.

Specifically, there was a great deal of interest amongst the majority of the music teachers, many of whom had a wealth of personal knowledge along with years of experience. Further discussions with these individuals may have produced additional, different ideas on what constitutes musical style and how children not only appear to view it, but, for example, how they use it within their own compositions. Certainly, more descriptions of the type given by Hargreaves (1999) as to the way in which children use stylistic elements in their own compositions, improvisations, drama and their playground games, would have yielded a great deal of interesting and valuable, yet very different supportive data of what the children actually knew of musical styles.

Secondly, although two previous studies have incorporated both pairwise and triadic test formats and experiment six of this thesis involved a further change in the testing methodology, all previous studies have employed an almost identical test format. Virtually all participants in style sensitivity tests have been presented with a similar musical task, a number of musical excerpts, an unknown test presenter and a response sheet to complete. Any inherent weaknesses in that test procedure which influence or distort the results, have possibly been present in all studies. Future research into this area would certainly benefit from a number of developments in the research methodology. The use of more longitudinal and qualitative methodologies run consecutively with the more quantitative studies reported here, could provide interesting and contrasting results. In the present thesis this could have been achieved both through the use of interested music teachers but also by involving several individual case studies involving fewer participants over a prolonged period of time with long term developments set within the context of a more detailed knowledge of the individuals musical experiences, both at home and at school. This would provide an interesting contrast to the more quantitative methodologies used in this and previous studies. Similarly, studies involving small groups of individuals which

explored the collective reactions to musical styles may provide a sound foundation to a developmental social psychology of music as outlined in chapter two.

On a more specific level, there were a number of issues which produced some further interesting questions.

Experiment Five attempted to explore further the issue of stylistic divergence. Two contrasting results were reported, namely the research by Gardner (1973) which had suggested that style sensitivity varied as a function of chronological distance between musical eras; whilst Tafuri et al. (1994), suggested that style sensitivity varied as a function of the distance a piece diverged stylistically from those styles which individuals had already internalised 'stylistic patternings'. The Italian study had suggested that individuals had internalised the 'stylistic patternings' of Baroque music and the music of Mozart specifically because these styles were far more prominent in the media. No empirical data was given to support this assumption and the interesting question arising from this experiment surrounded whether or not the individuals internalised those patternings because they were indeed more frequently used by the media, or whether individuals internalised those patternings first because of intrinsic features within the music itself.

Secondly, the interaction between presenter and participants with respect to age, culture and gender could prove to be a fruitful area for future research. For example, a significant participant / female experiment interaction was found in experiment two. Further information as to exactly what participants reacted to could account for some of the discrepancies arising between previous studies. Further studies surrounding 'who presents what music and to whom' would provide valuable information to music educators, musicians and a wide range of arts organisations.

Thirdly, the means of controlling the level of stylistic divergence between the individual musical excerpts is also a fruitful area for further research. Participants appear to attend to a wide range of varied musical elements in order to fulfil the task with which they have been presented and the experiments reported in this thesis suggest that minor variations in test material may produce different results. Some

means of more rigorously controlling stylistic divergence would benefit all areas of future research into style sensitivity.

It is vital to note that all previous studies involving both classical and poplar pairs had consistently reported participants performing significantly better on the popular rather than the classical pairs. Experiment five of this thesis suggested that the musical material utilised by any style sensitivity test, at least as operationalised here, can be crucial to the outcome and constitute a major influence on the participants' performance.

Fourthly, as all previous studies have shown, participants are capable of displaying a high level of stylistic sensitivity at a very early age. The employment of new test formats has shown that it is possible to explore style sensitivity in younger participants without increasing the linguistic element of the test. Further research, possibly involving the use of more longitudinal studies, and other methodologies, such as case studies of individuals and their developments, could shed far more light on the intermediate processes involved in the development of style sensitivity. Some participants are able to match musical styles by the age of three in spite of having little or no formal musical training. Further research into the individual backgrounds of these participants, involving comparisons of musical experiences in the home, would certainly be valuable. The particular study reported here has suggested that music education programs and training may not necessarily increase or contribute in any way major way towards the development of style sensitivity. Any research that increased the music educator's knowledge of the type of musical activities that both promote and support the development of style sensitivity would certainly be valuable.

Lastly, Hargreaves and Zimmerman (1992) remind us that in using the term 'development', that is, "overall changes in the patterning of behaviour that follow a regular and invariant sequence with age", a clear distinction needs to be drawn between:

" changes that are a product of *enculturation* and those that are a product of *training*. The former occur spontaneously in a given culture, without any conscious effort or direction, and the latter are the result of self conscious, directed efforts.....

We need to distinguish between the explanation of music learning and music teaching” (p.377).

If the development of style sensitivity can in fact be better described as a combination of skills, a clearer and more accurate understanding of the role played by cultural acquisition and music education in each individual skill, would be of great benefit.

The development of style sensitivity in music has been a much neglected area of research. Not only has there been a very small number of studies, but in many respects those studies themselves have been limited, for example Gardner’s study was carried out a very small population and style sensitivity was only one small part of much broader studies carried out by Castell (1983) and Hargreaves and North (1997). In essence, very little was known about this subject with all previous studies being confined to an almost identical test methodology. Many questions existed about the exact nature of the relationship between the participant and the social context of the music, their music education programme, the test presenter, the test format and the artefactual elements of the music used in the samples. In the next section, a model of development is proposed. In producing this proposed model a number of things were taken into account. Firstly, Hargreaves (1986) reminded us that early work in developmental psychology tended to be ‘descriptive’ rather than ‘explanatory’. Similarly, Sloboda (1985) also reminded us that knowing what a skill entails, is very different from executing it. In producing the model, some effort was made to move beyond a view and description of style sensitivity as a single unitary skill and instead to suggest a possible mental process by which decisions surrounding issues of musical style are made by individuals. Secondly, some effort was made to produce a model which could accurately account for the results obtained not only in the present research, but also in previous studies.

In 1973, Gardner noted that exploration of the development of style sensitivity to musical style was an interesting area for research. The resulting 25 years have seen only a limited number of studies, all of which have operated within an almost identical paradigm, and which have made a number of quite significant claims as to why subjects should perform differently on the classical and the popular styles. Little effort has been made however, to establish the validity of the instrument of measurement. The resulting data from experiments one to six of this thesis have

suggested that style sensitivity tests, at least as operationalised here, are subject to a wide variety of spurious influences. Participant's performance may be affected by the musical sample used, the way the test is presented, their motivation to perform in the test and the context within which the test is set. Results from all previous studies should therefore be viewed in this light.

11.4 Discussion: Style sensitivity to stylistic discrimination - a proposed model of development.

In his original study, Gardner (1973) defined style sensitivity as the ability to classify whether two extracts of music came from the same or from different compositions. Gardner also suggested that the activity could manifest deeper cognitive processes in action which involves the monitoring of numerous aspects of a stimulus. However, one might question whether or not the paradigm he devised is too narrow to measure 'complex cognitive processes'. Furthermore, although subsequent studies have added to and extended the original Gardner study, all have used an almost identical test methodology. Any limitations as to the reliability and the validity of the original instrument may well therefore be present in all subsequent studies.

It could be argued that the reliability of the test instrument is quite high since all studies have come to broadly similar developmental conclusions. That is, whilst there have been subtle differences between studies, there have also been a number of recurring phenomena, e.g. the fact that the highest means always occur for the younger participants and not the older ones. Whilst it would appear that the instrument of measurement does demonstrate some reliability, it is necessary to further question the validity of that instrument, namely what exactly is it measuring reliably?

Certainly, the experiments carried out in this thesis have suggested that participants' performance on the style sensitivity task can vary significantly as a function of the test presenter, the musical samples used and the way in which the test is introduced to the participants. The claims made by Castell (1983) for example, that younger participants performed more accurately because they were more 'open eared', should therefore be viewed in the light of other influences on the test performance. Similarly,

the claims made by Addessi et al., that their participants performed better on musical examples from the classical period because they had internalised the ‘patterning’ from the music of this period more thoroughly, should be viewed in the light of experiment five of this thesis, in which it was possible to vary participants’ performance on the classical pairs through manipulations of the musical examples.

The ecological validity of an experiment questions how closely the results can be generalised to everyday settings. Carlsmith et al. (1976) suggests that the term “mundane realism” can be used for experiments which are close to real life, but that the term “experimental realism” is more appropriate for those situations in which the experimental task, although artificial, is so engaging and demanding of attention that any artificiality is compensated for. In terms of Gardner’s paradigm, one must question whether or not the task itself has ecological validity, and therefore how well it can be generalised to other situations. Do participants classify music along the lines required by the test in everyday life? The answer is probably that they do not.

However, the consistent finding that style sensitivity does appear to decrease in the early teenage years before increasing once again, does require some further comment. There are two possible reasons for this decrease; first, that it may be a genuine developmental phenomenon. Second, however it could be that some other influence or process is being measured by the test format. That is, the test instrument may be sufficiently sensitive to measure style sensitivity in younger subjects, whilst older subjects incorporate further mental processes into their decision-making within the test situation. Therefore although style sensitivity does increase, (or to reach a certain level in young children and then to remain at that level), the test ceases to discriminate, and becomes sensitive to the influence of other mental processes.

Perhaps younger children possess sufficient levels of sensitivity to musical style to enable them to identify any differences between musical samples. As Sloboda (1985) states; they have the ability “to perceive and remember invariant aspects of otherwise different patterns” (p.210) in the musical pieces. However, they may not necessarily be classifying the music into various styles. Older children may well have the same, or indeed an increased sensitivity to identify the differences and similarities between the musical examples, but are explicitly attempting to classify the music into styles

which could be seen as a more logical, linguistic activity. In other words, the younger and older subjects may be carrying out very different mental activities within the test: the younger subjects simply identifying differences whereas the older ones may be classifying music into distinct styles, which is dependent on a set of higher order mental skills.

Further to this, previous studies have tended to use varying terminologies, and sometimes to use them interchangeably. Style sensitivity, stylistic competence and stylistic discrimination have all been used in reporting the work in this area, but not all have given an adequate description of the precise meaning and the way each term is being used. In the light of the previous comment, that younger children may be able to identify differences, whereas older subjects are attempting to classify, perhaps more specific definitions are now required. These might be as follows:

Stylistic Sensitivity: The ability to perceive and remember invariant aspects of otherwise different patterns in musical pieces.

Stylistic Discrimination: The ability to classify various musical excerpts into musical styles.

Previous studies have tended to view style sensitivity as a single developing unitary skill, and have suggested that the differences in results between studies occur because different social influences may impact upon the efficiency with which the individual may utilise that developing skill on the test itself. It therefore seems sensible to ask exactly what the test may be measuring in older subjects. Hargreaves (1986) suggested that early work in developmental psychology tended to be ‘descriptive’ rather than ‘explanatory’. Similarly, Sloboda (1985) also proposed that knowing what a skill entails is very different from executing it. In the next section, I should like to propose a tentative model which spells out some of the influences which may influence the decision making process which takes place within the style sensitivity tests employed in this thesis.

In producing the model, some effort was made to move beyond a view and description of style sensitivity as a single unitary skill, and instead to map out a

mental process by which decisions surrounding issues of musical style are made by individuals. Secondly, some effort was made to produce a model which may account for the results obtained not only in the present research, but also in previous studies. Lastly, it is hoped that the model could provide a theoretical framework for further research into this area.

It was noted in an earlier chapter that Koroscik (1997) had proposed that ‘understanding’ was arrived at through three individual but interrelated facets of cognition. She refers to the first of these as the individual knowledge base and to the second as the individual knowledge strategy, which is the strategic process the individual takes to apply knowledge or skills gained in one area to a new and different area.

The third facet identified by Koroscik is similar to Swanwick’s (1994) ‘attitudinal’ level of knowledge; namely the level of motivation. This is the motivation to acquire more knowledge, and correspondingly to apply transfer strategies to new knowledge. Koroscik suggested that success in understanding a work of art is dependent on the individual successfully and equally employing all of her three cognitive strands; namely factual information, transfer skills and motivation. With specific reference to style sensitivity, this third facet, i.e. motivation, should be most clearly affected by the social influences. Studies into style sensitivity might thus be viewed and interpreted within the following theoretical framework.

1. Musical knowledges. This includes the acquisition of various aspects of musical knowledge and skill, and experiences / influences which affect this e.g. music education.
2. Transfer strategies. This includes studies of the process whereby individuals select and apply musical knowledges to the musical cues they are able to discern. (Experiment 6)
3. Motivational or attitudinal. This includes studies of factors and social influences affecting the level of motivation to process music and to develop new musical understandings. This would include social psychological perspective. (Experiments 1, 2 and 4)

Style sensitivity therefore could be described as follows. The musical input is received and processed at one of the lower-order levels of analysis proposed earlier by Meyer and Nattiez. Style sensitivity could be said to take place at the level of “Universal features” and/ or “Reference system” in Nattiez’ terminology. Participants have acquired sufficient knowledge of their cultural reference system in order to identify invariant aspects within the two musical samples.

However, the result of this processing of information at this level may be further influenced by the additional logical task required by the test namely, the need to record whether or not the two pieces of music are from the same or a different piece of music. Indeed the final response recorded on the test response sheet may well be influenced by a number of other elements such as the understanding of the language / requirement of the task and the social and other factors affecting the motivation to perform the task well.

Stylistic discrimination may be a much more complicated process, however, aspects of which may complicate or mask style sensitivity as such: it may take place at a much deeper level. To revert to Nattiez’ terminology, processing may be taking place at the level of ‘genre’ or ‘composer.’ In some instances, it may even take place at the level of ‘period’ or ‘work’. Operating at this level would involve far greater transfer skills and far more motivation, and it may well result in the test methodology ceasing to measure sensitivity to musical style differences, and instead to highlight another developmental phenomena. For example, older participants may be aware of the need to apply their musical knowledge to the musical stimulus, but may lack the transfer skills to apply their knowledge correctly. Similarly, they may lack the motivational skills to perform the more complex task.

11.5 Novice to Expert

Two further issues deserve some comment. Firstly, as explained in chapter three, White (1997) challenged all developmental theories by questioning the idea of an end state of development. White suggested that mental development, unlike physical development, might possibly proceed with ‘endless growth in certain directions’. Certainly, it seems at least to be a possibility that the development of the domain of

music, that is musical knowledge, can continue to grow throughout life. Secondly, Koroscik (1997) speaks of development not as a series of stages, but as movement along a continuum from the state of 'novice' at one pole to the state of 'expert' at the opposite pole. As Koroscik points out, the term 'novice' not only applies to most children, but also to any person, regardless of age, whose understandings are underdeveloped within a given area. Similarly, Gardner (1973) speaks of movement from the 'novice' to the 'connoisseur'. 'Connoisseurs' or 'experts' will incorporate everything into their analytical decision making, from the 'age and condition of canvass to the style of brush stroke' (p.).

In terms of sensitivity to musical style, the movement along the continuum from novice to expert may include the acquisition of further musical knowledges. However, it may also involve the skill of knowing which level of analysis to apply and when to apply it, as appropriate to the task required. That is, older subjects may have the motivation to complete the task; they may have the musical knowledges to complete the task on a simple level. However, they may not possess either the necessary transfer skills or the ability to correctly select the most appropriate level on which to make their analysis. This could explain, for example, the difference in scores between the older participants and the young adults in Castell's study. Both groups had well developed sensitivity and both groups possessed high levels of musical knowledge. However, the adult group also possessed the knowledge / experience to know when and how to select the appropriate level of analysis and how to transfer that knowledge: this was demonstrated in experiment 5, for example, in which the levels of contrast between the classical and popular pieces were manipulated. The presence of prominent elements within the music perhaps made it easier for older subjects to revert to a less complex level of analysis, and their performance correspondingly improved.

However, the ability of Castell's adult participants to correctly identify all the musical examples as coming from the same or different pieces should not necessarily be labelled as 'expert'. Indeed, it is probable that these participants performed well because the knowledge required in processing and classifying the musical pairs fell into the levels of analysis for which they possessed the knowledge. That is, it was possible to process the differences between the specific musical excerpts used because

no musical knowledge was required beyond the level of 'genre', in Nattiez' terminology. For example, the differences between the exact excerpts selected to represent the classical period and the romantic period could be processed with limited knowledge within the lower levels of knowledge i.e. 'universal features', 'reference system' and 'genre'. Had two excerpts been presented which required more detailed musical knowledge from the higher levels of 'composer', 'period' or 'work', for example a piece by Saint-Saens and a piece by Grieg or Pearl Jam and Metallica, then the adults may have performed less well.

In other words, to become an 'expert' requires having the ability to process musical artefacts on higher levels of Nattiez' inverted pyramid. An advanced level of development along the continua of stylistic competence would not be the mature adult with the ability to simply identify two excerpts as both being Classical, Baroque, Ska or Grunge, for example, but rather the expert / connoisseur who identifies which notes of the Requiem were composed by Mozart and those written by his pupil Sussmeyer.

If style sensitivity is defined as the ability to identify differences in musical examples and stylistic discrimination is the ability to classify music on a variety of levels, perhaps stylistic competence might perhaps be described as occurring within a narrow musical field when the individual has the appropriate knowledge base to attempt analysis on all Nattiez' levels and has the motivation to complete the task. This level of functioning also involves the skills to transfer knowledge to that which they are hearing. This involves correctly selecting and utilising the appropriate analytical level at pertinent points within the music.

11.6 Summary

Style sensitivity is a fascinating and relatively under researched area of music psychology and it is hoped that this study will go some way towards promoting further interest in this area. One of the stated aims of this study was that it should augment and further the knowledge and understanding of the way in which concepts of musical style develop in individuals and it is appropriate to question whether or not this has been achieved. I would like to suggest that this study has contributed a number of things to the overall study and knowledge of the development of stylistic sensitivity.

Firstly, some attempt has been made to identify, distinguish and propose a more precise terminology. Previously, terms such as style sensitivity and stylistic competence have been used somewhat interchangeably. I have argued here that these essential terms may in fact be describing very different mental processes and some degree of agreement on the precise use of these terms would be of help to future studies.

Secondly, in one sense, the concept of stylistic sensitivity has previously been treated as a single event. It is hoped that this study has gone some way towards promoting the fact that measurements of stylistic awareness do not represent a single growing musical ability, the developmental path of which is linear. Measurements of stylistic awareness can be more accurately seen as a multidimensional description of a combination of abilities, skills and knowledge. Each individual component, and the way it interacts within the total process is worthy of further study.

Thirdly, it is hoped that this study has gone some way towards consolidating the results of previous studies. Some effort has been made to unpick the tangle of influences that may possibly have affected the results of previous studies. Hargreaves and North (1999) pointed out:

“...it is probably impossible to disentangle the cognitive aspects of responses to musical styles from their social and affective components, and the interaction between all three probably gives rise to inconsistent age trends in different studies” (p. 200)

The findings of the experiments reported in this thesis suggest that a large number of additional variables relating to the test, test material, test presenter and test setting may also be added to the list.

There is evidence that very young children have in their possession a large number of developed musical capacities and an endless potential to develop them. Yet it is often the case that many adults only achieve a fraction of their musical potential. In many cases, whatever potential this has occurred independently of music education programs. Perhaps the time has come to now re evaluate our musical education provision. Perhaps we have been too guilty of a ‘top down approach’; of expressing

our wishes as to what it would be desirable for children to be able to do by a certain age and designing a music education program in order to fulfil those objectives. Perhaps we now need to be discovering and describing the immense musical abilities that children possess at a very young age, and devising musical experiences which build upon and develop those abilities more fully.

Discovering, identifying and contributing towards a model of developmental stages in musical activities are exciting, rewarding, interesting and necessary tasks for the musician or music educator. Examining how children process, understand and attribute meaning to musical sounds must be a key area for all educators who want to address the issue of what, when and how to present and use meaningful music to their pupils. The problem goes beyond simply establishing what music children find important. Research needs to establish what musical elements children perceive and when they perceive them; how they create some form of order from the diverse range of musical elements they hear; and how they draw together the relationships between the various musical aspects and create musical judgements. The task of the music educator is to provide teaching styles and teaching materials that will assist this constantly developing process. It is hoped that this study has gone some small way towards that goal.

Appendix One

Experiments one to four: musical samples.

Popular Excerpts.

Grunge: Sugar

Thrash: Nuclear Assault

The Act We Act
Changes
Man on the Moon

Critical Mass
Game Over
Stranded in Hell

Indie: Echo and the Bunnymen

Blues: B.B.King

The Puppet
Do it Clean
Show of Strength

Meet my Happiness
You Put it on Me
Don't Answer the Door

Classical Excerpts.

Baroque
J.S.Bach:

Classical
W.A.Mozart

"Courente", French Suite n.4 (Eflat maj.)
"Courente", French Suite n.5 (Gmaj.)
"Courente", French Suite n.6 (E maj.)

"Allegro assai" Piano Sonata F maj.
"Presto", Piano Sonata in G maj.
"Allegro", Piano Sonata in D maj.

Romantic

Modern

F.Chopin

S.Prokofiev

"Scherzo" Piano Sonata in Eflat maj.
"Valzer" in D flat maj.
"Preludio" in G maj.

"Allegro strepitoso" in C maj.
"Allegro con brio" in C maj.
"Allegro con brio" in C Maj.

Appendix Two

Experiment five: musical samples

Baroque:

1. Concerto Grosso Nos.1 G.F.Handel
2. Concerto Grosso Nos.2 G.F.Handel
3. Brandenburg Concerto Nos.3 J.S.Bach

Classical:

1. Eine Kleine Nacht Musik. 2nd. Movement. W.A.Mozart
2. Serenade for Wind Instruments. 1st. movement. W.A.Mozart

Romantic:

1. Symphony Nos. 2 3rd. movement. S.Rachmaninov
2. Tristan and Isolde overture. R.Wagner

Modern:

1. Ritual Dances. M.Tippet
2. Symphony Nos.7. 3rd. movement. D.Shostakovitch

Popular Styles

Ska: No Doubt

1. Happy Now
2. End it on this
3. Different People
4. Hey you

Jazz: Frank Zappa

1. Son of Mister Green Genes
2. Little Umbrellas
3. It Must Be A Camel
4. Gumbo Variations

Grunge: Pearl Jam

1. Once
 2. Even Flow
- ##### Grunge: Sound Garden
3. Outshined

Thrash: Mettalica

1. Devil's Dance
2. Carpe Diem Baby
3. Prince Charming
4. Attitude

Appendix Three.

Experiment six: musical samples

Popular Examples

New Age: David Akenfield
Moonflower

Vanessa May (Metal)
Hocus Pocus

Ultravox (Techno)

Yanni (Electric Jazz)

One Fine Day. (Remix)

Classical Excerpts.

Baroque
J.S.Bach:

“Courente”, French Suite n.4 (Eflat maj.)
“Courente”, French Suite n.5 (Gmaj.)
“Courente”, French Suite n.6 (E maj.)

Classical
W.A.Mozart

“Allegro assai” Piano Sonata F maj.
“Presto”, Piano Sonata in G maj.
“Allegro”, Piano Sonata in D maj.

Romantic

F.Chopin

“Scherzo” Piano Sonata in Eflat maj.
“Valzer” in D flat maj.
“Preludio” in G maj.

Modern

S.Prokofiev

“Allegro strpitso” in C maj.
“Allegro con brio” in C maj.
“Allegro con brio” in C Maj.

Bibliography

- Abbs, P. (1987). *Living powers: the arts in education*. Brighton: Falmer Press.
- Abeles, H.F. (1980). Responses to music, in D. Hodges (ed.). *Handbook in music psychology*. Lawrence, Kansas: National Association for Music Therapy.
- Addessi, A.R.; Baroni, M.; Luzzi, C., and Tafuri, J. (1995). *The development of musical stylistic competence in children*. Council for Research in Music Education Bulletin, **127**, 8-15.
- Addessi, A.R.; Luzzi, C. and Tafuri, J. (1996). *La competenza stilistica musicale dagli 8 ai 14 anni*, in J. Tafuri (ed.), *La comprensione degli stili musicali*. Quaderni della SIEM, **10**, 59 –80.
- Adorno, T. (1994). *Quasi una fantasia: essays on modern music*. London: Verso.
- Aiello, R. (1994). Can listening to music be experimentally studied? in R. Aiello and J. Sloboda (Eds.). (1994). *Musical perceptions*. Oxford: Oxford University Press.
- Aiello, R. & Sloboda, J. (Eds.). (1994). *Musical perceptions*. Oxford: Oxford University Press.
- Aiello, R.; Tanaka, J. and Winborne, W. (1990). Listening to Mozart: perceptual differences among musicians. Journal of Music Theory Pedagogy, **4**(2), 269-93.
- Allen, S.A.; Dubanoski, R.A. and Stevenson, H.W. (1966). Children's performance as a function of race of E, race of S, and type of verbal reinforcement. Journal of Experimental Psychology, **4**, 248-256.
- Anastasi, A. (1988). *Context of psychological testing*. New York: Macmillan Publishing Company.
- Areni, C.S. and Kim, D. (1993). The influence of background music on shopping behaviour: classical versus top-forty music in a wine store. Advances in Consumer Research, **20**, 336-40.
- Asch, S.E. (1948). The doctrine of suggestion, prestige and imitation in social psychology. Psychological Review, **55**, 250-276.
- Asch, S.E. (1956). Studies of independence and conformity. A minority of one against a unanimous majority. Psychological Monographs, **70**, No.416.
- Attali, J. (1985). *Noise: The political economy of music*. Minneapolis: University of Minnesota Press.

Bamberger, J. (1979). *The art of listening: developing musical perception*. (4th edition). New York: Harper and Row.

Bamberger, J. (1980). Cognitive structures in the apprehension and description of simple rhythms. Archives de Psychology. XLVLII, 171-199.

Bamberger, J. (1994). Learning to hear in a new way. in R.Aiello and J. Sloboda. (Eds.). (1994). *Musical perceptions*. Oxford: Oxford University Press.

Bandura, A. (1971). *Psychological modelling*. Chicago: Adline-Atherton Inc.

Bantock, G.H. (1968). *Culture, industrialisation and education*. London: Routledge and Kegan Paul.

Bartlett, J.C. & Dowling, W.J. (1980). The recognition of transposed melodies: a key-distance effect in developmental perspective. Journal of experimental Psychology: Human Perception and Performance, 6, 501-15.

Baumann, V.H. (1960). *Teenage music preferences*. Journal of Research in Music Education. 8, (2), 75-84.

Baumgarten, A. (1936). (first published 1735). Reflections on poetry. In Aesthetica. B.Croce. (ed.). Rome: Bari.

Beardsley, M.C. (1966). *Aesthetics: from classical Greece to the present*. London: University of Alabama Press.

Beardsley, M.C. (1981). *Aesthetics: problems in the philosophy of criticism*. Cambridge: Hackett Publishing Company, Inc.

Bergen, A.; McManis, D.L. & Melchert, P.A. (1971). *Effects of social and token reinforcement on WISC block design performance*. Perceptual and Motor Skills. 32, 871-880.

Berlyne, D.E. (1972). Experimental aesthetics. in *New Horizons in Psychology* 2. P.C.Dodwell. (ed.). Penguin: Harmondsworth.

Bigand, E. (1990). Abstraction of two forms of underlying structure in a tonal melody. Psychology of Music, 18, 45-59.

Binder, A.; McConnell, D. & Sjöholm, N. (1957). Verbal conditioning as a function of experimenter characteristics. Journal of Abnormal Social Psychology. 55, 309-314.

Bittner, A.C. & Shinedling, M.M. (1968). A methodological investigation of Piaget's concept of conservation of substance. Genetic Psychology Monographs. 77, 135-165.

Bjurstrom, E. & Wennhall, J. (1991). *Ungdomar och music. Arsbok om ungdom*. Stockholm: Statens Ungdomsrab.

Blacking, J. (1973). *How musical is man?* Seattle: University of Washington Press.

Boselie, F. (1996). *Prototypicality revisited: A rejoinder to Hekkert and Snelders*. Empirical Studies of the Arts. **14** (1), 99-104.

Bradbury P.J.; Wright, S.D.; Walker, C.E. & Ross, J.M. (1975). Performance on the WISC as a function of E, sex of S, and age of S. Journal of Psychology **90**, 51-55.

Bransford, J.; Sherwood, R.; Vye, N. & Rieser, J. (1986). Teaching thinking and problem-solving: Research foundations. America Psychologist, **41** (10), 1078-1089.

Brant, L., Marshall, P. & Roark, B. (1995). On the development of prototypes and preferences. Empirical Studies of the Arts. **13** (2). 161-170.

Bridgeman, B. (1974). Effects of test score feedback on immediately subsequent test performance. Journal of Educational Psychology. **66**,(1), 62-66.

Brittin, R.V. & Sheldon, D.A. (1995). Comparing continuous versus static measurements in music listeners' preferences. Journal of Research in Music Education. **43**, (1).

Brittin, R.V. (2000). Children's preference for sequenced accompaniments: The influence of style and perceived tempo. Journal of Research in Music Education. **48**,(3). 237-248.

Bronowski, V. (1973). *The ascent of man*. London: BBC Publications.

Brown, G. (1983). *Piaget's theory and educational psychology*. in S. Modgil., C. Modgil., & G. Brown (eds.). *Jean Piaget: An interdisciplinary critique*. London: Routledge and Kegan Paul.

Brown, T. & Weiss, L. (1987). Structures, procedures, and affectivity. Archives de Psychologie. **55**, 59-94.

Bryant, P.E.(1974). *Perception and understanding in young children*. London: Methuen.

Bryant, P.E. (1982). Piaget: Issues and experiments. British Journal of Psychology. **73**. (2).

Budd, M. (1985). *Music and the emotions: the philosophical theories*. London: Routledge.

Bullock, W.J. (1973). A Review of measures of music-aesthetic attitude. Journal of Research in Music Education, **21**, 331-344.

- Carroll, W. (1910). The unfolding personality. Unpublished lecture notes.
- Carroll, W. (1914). *The unfolding personality*. London: Sherratt and Hughes.
- Carlsmith, A. (1976) in Coolican, H. (1999).
- Castell, K. (1982). Children's Sensitivity to Stylistic Differences in 'Classical' and 'Popular' Musical Pairs. Psychology of Music. Special Issue, 22-25.
- Castell, K. (1983). Responses to music. Unpublished Ph.D. thesis. University of Leicester.
- Chang, W. & Trehub, S. E., (1977). Auditory processing of relational information by young infants. Journal of Experimental Child Psychology, **24**, 324-31.
- Chapman, A.J. & Williams, A.R. (1976). *Prestige effects and aesthetic experiences: adolescents' reactions to music*. British Journal of Social and Clinical Psychology, **15**, 61-72.
- Cieutat, V.J. & Flick, G.L. (1967). *Examiner differences among Stanford-Binet items*. Psychological Reports, **21**, 613-622.
- Clarke, E.F. (1989). Issues in language and music. Contemporary Music Review, **4**, 9-22.
- Clarke, E.F. & Krumhansl, C. (1990). Perceiving musical time. Music Perception, **7**, 213-51.
- Cohen, M.R. (1944). *A preface to logic*. New York: Henry Holt and Co.
- Cohen, L. & Manion, L. (1980). *Research methods in education*. (4th edition). London: Routledge.
- Colwell, R. (Ed.). (1992). *Handbook of research on music teaching and learning*. New York: Schirmer Books.
- Cook, N. (1987). The perception of large- scale tonal closure. Music Perception, **5**, 197-205.
- Cook, N. (1990). *Music, imagination and culture*. Oxford: Oxford University Press.
- Cook, N. (1994). Perception: a perspective from music theory. in R. Aiello and J. Sloboda. (Eds.). (1994). *Musical perceptions*. Oxford: Oxford University Press.
- Cooke, D. (1959). *The language of music*. Oxford: Oxford University Press.

- Cooke, P. (1978). Music-learning in traditional societies. (29-34) in R. Leach & R. Palmer. *Folk Music in schools*. Cambridge: Cambridge University Press.
- Coolican, H. (1999). *Research Methods and Statistics in Psychology*. London: Hodder and Stoughton.
- Cooper, R.P. & Aslin, R.N. (1989). The language environment of the young infant: implications for early perceptual development. *Canadian Journal of Psychology*. **43**, 247-65.
- Cortazzi, M. (1993). *Narrative analysis*. London: Falmer Press.
- Cotterell, J. (1996). *Social networks and social influences in adolescents*. New York: Routledge.
- Covington, M.V. & Omelich, C.L. (1979). *Effort: the double - edged sword in school achievement*. *Journal of Educational Psychology*. **71**, (2). 169-182.
- Cox, G. (1990). The legacy of folk-song: the influence of Cecil Sharp on music education. *British Journal of Music Education*. **7**, (2).
- Cox, G. (1993). *History of music education in England. 1872-1928*. Hants: Scolar Press.
- Cronbach, L. (1990). *Essentials of psychological testing*. (5th edition). New York: Harper Collins.
- Cross, I. (2001). Music, mind and evolution. *Psychology of Music* **29**, (1) 95-102.
- Crozier, J.W. & Chapman, A. (1981). *Aesthetic preferences: prestige and social class*. in D. O'Hare. (ed.). *Psychology and The Arts*. Brighton: Harvester Press.
- Crummer, G.; Hantz, E.; Chuang, S. & Walton, J. (1988). Neural basis for music cognition: initial experimental findings. *Psychomusicology*. **7**, 117-126.
- Csikszentmihalyi, M. & Robinson, R.E. (1990). *The art of seeing: an interpretation of the aesthetic encounter*. Malibu, Ca.: The J. Paul Getty Trust.
- Cuddy, L. & Cohen. A. (1976). Recognition of transposed melodic sequences. *Quarterly Journal of Experimental Psychology*. **28**, 255-270.
- Davidson, J.W.; Howe, M.J.A. & Sloboda. J. (1997). Environmental factors in the development of musical performance skill over the life span. in D.J.Hargreaves and A.C. North. (eds.). *The Social Psychology of Music*. Oxford: Oxford University Press.
- Davies, J.B. (1971). *New Tests of Musical Aptitude*. *British Journal of Psychology*. **62**, (4). 557-565.

- Davies, J.B. (1978). *The psychology of music*. London: Hutchinson.
- Davis, W.E. (1969). *Examiner differences, prior failure and subjects' arithmetic scores*. Journal of Clinical Psychology. **25**, 178-180.
- DeCaspar, A.J. & Fifer, W.P. (1980). On human bonding: newborns prefer their mothers' voices. Science, **208**, 1174-6.
- DeCaspar, A.J. & Spence, M.J. (1986). Newborns prefer a familiar story over an unfamiliar one. Infant Behaviour and Development, **5**, 261-76.
- Deliege, I. (1989). A perceptual approach to contemporary musical forms. Contemporary Music Review, **4**, 213-30.
- Deliege, I. (1996). Cue abstraction as a component of categorisation process in music listening. Psychology Of Music. **24**, 131-156.
- Deliege, I. & El Hamadi, A. (1990). Mechanisms of cue extraction in musical groupings: a study of perception of Sequenza IV for solo viola by Luciano Berio. Psychology of Music, **18**(1), 18-44.
- Deliege, I. & Sloboda, J. (eds.). (1996). *Musical beginnings. Origins and development of musical competence*. Oxford: Oxford University Press.
- DeNora, T. (1995). *Beethoven and the construction of genius*. Berkley: University of California Press.
- DeNora, T. (2000). *Music in everyday life*. Cambridge: Cambridge University Press.
- D.E.S. (1992). *Music in the National Curriculum* (England). London: HMSO
- Deutsch, D. (1982). The processing of pitch combinations. in D.Deutsch. (ed.). (1982). *The psychology of music*. New York: Academic Press.
- DeWitt, L. & Samuel, A. (1990). The role of knowledge based expectations in music perception: evidence from musical restoration. Journal of Experimental Psychology: General. **119**, (2), 123-144.
- Dickstein, L.S. & Ayers, J. (1973). *Effect of an incentive upon intelligence test performance*. Psychological Reports. **33**. 127-130.
- Dickstein, L.S. & Kephart, J.L. (1972). *Effect of explicit examiner expectancy upon WAIS performance*. Psychological Reports. **30**, 207-212.

Diener, C.I. & Dweck, C.S. (1978). *An analysis of learned helplessness: continuous changes in performance, strategy, and achievement cognitions following failure*. Journal of Personality and Social Psychology. **36**,(5), 451-462.

Doise, W.; Mugny, G. & Perret-Clermont, A. -N. (1975). *Social interaction and the development of cognitive operations*. European Journal of Social Psychology, **5**, 367-83.

Doise, W. & Mugny, G. (1984). *The social development of the intellect*. Oxford: Pergamon.

Doise, W. (1986). *Levels of explanation in social psychology*. Cambridge: Cambridge University Press.

Doise, W. (1987). *Groups and individuals: explanations in social psychology*. Cambridge: Cambridge University Press.

Dolgin, K.G. & Adelson, E.H. (1990). Age changes in the ability to interpret affect in sung and instrumentally presented melodies. Psychology of Music. **18**, 87-98.

Donaldson, M. (1978). *Children's minds*. London: Fontana / Open Books.

Dorow, L.G. (1977). *The effect of teacher approval / disapproval ratios on student music selection and concert attentiveness*. Journal of Research in Music Education. **25**, 32-40.

Dowling, W.J. (1973). The perception of interleaved melodies. Cognitive Psychology. **5**, 322-337.

Dowling, W.J. (1978). Scale and contour: two components of a theory of memory for melodies. Psychological Review **85**, 341-354.

Dowling, W.J. (1982). Melodic information processing and its development. In D.Deutsch (ed.). *The Psychology of Music*. New York: Academic Press.

Dowling, W.J. (1989). Simplicity and complexity in music and cognition. Contemporary Music Review, **4**, 247-53.

Dowling, W.J. & Fujitani, D.S. (1971). Contour, interval and pitch recognition in memory for melodies. Journal of the Acoustical Society. **49**, 524-531.

Drake, R.M. (1939). *Factorial analysis of music tests by the Spearman-Tetrad-Difference technique*. Journal of Musicology, **1**, (1) 6-16.

Duerksen, G.L. (1972). *Some effects of expectation on evaluation of recorded musical performances*. Journal of Research in Music Education. **20**, 268-272.

Dunn, R.E. (1997). *Creative thinking and music listening*. Journal of Research Studies in Music Education. **8**.

Durkheim, E. (1974). *Sociology and philosophy*. New York: Free Press.

Durkin, K. (1995). *Developmental social psychology: from infancy to old age*. Oxford: Blackwell.

Durkin, K. (1996). Developmental social psychology. in M. Hewstone., W. Stroebe & G.M. Stephenson. (eds.). *Introduction to social psychology*. 2nd. edition. Oxford: Blackwell.

Dweck, C.S. & Leggett, E.L. (1988). *A social - cognitive approach to motivation and personality*. Psychological Review, **95**, 256 - 273.

Dyer. H.S. (1973). *Recycling the problems of testing*. Procedures of the 1972 invitational conference on testing Problems. Educational testing service. 85-95.

Ecker, D. (1963). The artistic process as qualitative problem – solving. Journal of Aesthetics and Art Criticism. **21**, 283-290.

Elkind, D. (1976). *Child development and education: A piagetian perspective*. New York: Oxford University Press.

Emler, N.; Ohana, J. & Moscovici, S. (1987). Children's beliefs about institutional rules: a cross national study of representations of the teacher's role. British Journal of Educational Psychology. **57**, 26-37.

Ericsson, K.A.; Kampe, R.T. & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. Psychological Review, **100**, 363-406.

Exner. J.E. (1966). *Variations in WISC performance as influenced by differences in pre test rapport*. Journal of General Psychology. **74**, 299 – 306.

Farnsworth, P.R. (1969). *The social psychology of music*. (2nd Edition). Ames: Iowa State University Press.

Farnsworth, P.R. & Beaumont (1929). Suggestions in pictures. Journal of General Psychology. **2**, 362-6.

Fechner, G.T. (1876). *Vorschule der aesthetic*. Leipzig: Breitkopf and Hartel.

Ferguson, D.C. & Buss, A.H. (1960). Operant conditioning of hostile verbs in relation to experimenter and subject characteristics. Journal of Consulting Psychology. **24**, 324-327.

Fiske, H. (1992). Structure of cognition and music decision-making. in R.Colwell (ed.), *Handbook of Research on Music Teaching and Learning*. New York: Schirmer Books.

Flowers, P.J. (1983). The effect of instruction in vocabulary and listening on non musicians' descriptions of changes in music. Journal of Research in Music Education. **3**, 179-189.

Flowers, P.J. (1984). Attention to elements of music and effect of instruction in vocabulary on written descriptions of music by children and undergraduates. Psychology of Music. **12**, 17-24.

Flowers, P.J. (1988). The effects of teaching and learning experiences, tempo, and mode on undergraduates' and children's symphonic music preferences. Journal of Research in Music Education. **1**, 19-34.

Flowers, P.J. (1990). Listening: The key to describing music. Music Education Journal. **77**, 21-23.

Fodor, J.A. & Bever, T.G. (1965). The psychological reality of linguistic segments. Journal of Verbal Learning and Verbal Behaviour, **4**, 414-421.

Forsyth, J.L. (1975). The effect of teacher approval, disapproval, and errors on student attentiveness: music versus classroom teachers. in C.K. Madsen., R.D.Greer. & C.H.Madsen, Jr. (eds.) Research in Music Behaviour. New York: Teachers College Press.

Fraisse, P. (1982). Rhythm and Tempo. in D.Deutsch, (ed.). (1982). *The psychology of music*. New York: Academic Press.

Frances, R. (1967). Communication persuasive et communication esthetique. Journal de Psychologie, **4**, 415 – 30.

Freud, S. (1938). Limitations of psychoanalysis: S.Freud, "Infantile Sexuality", in A.A. Brill (ed.). *Basic writings of Sigmund Freud*. New York: Random House.

Freud, S. (1972). *Two short accounts of psychoanalysis* London: Pelican.

Friedman, N. (1964). The psychological experiment as a social interaction. Unpublished Ph.D Thesis. Harvard University.

Forsyth, J.L. (1975). The effect of teacher approval, disapproval, and errors on student attentiveness: music versus classroom teachers. in *Research In Music Behaviour: Modifying Music Behaviour in the Classroom*. 1975. C.K. Madsen., D.R.Greer., and C.H. Madsen., Jr. New York: Teachers College Press.

Furth, H.G. (1980). *The world of grown- ups: children's conceptions of society*. New York: Elsevier.

Gabrielsson, A. (1993). *Emotion and Music*. ESCOM. Newsletter 4.

Gahagan, J.P.; Long, & Horai, J. (1969). Race of experimenter and reactions to threats by black pre adolescents. in J. Sattler. *Racial 'Experimenter Effects' in Experimentation, Testing, Interviewing, and Psychotherapy*. Psychological Bulletin. 73, (2), 137-160.

Gallagher, J. & Reid, D. (1981). *The learning theory of Piaget and Inhelder*. Monterey, Ca.: Brooks-Cole.

Galton, M. & Delamont, S. (1985). Speaking with forked tongue? Two styles of observation in the ORACLE project. in R.G. Burgess (ed.). *Field Methods in the study of Education*. London: Falmer Press.

Gardner, H. (1970). Children's sensitivity to painting styles. Child Development. 41, 813-21.

Gardner, H. & Gardner, J. (1970). Developmental trends in sensitivity to painting style and subject matter. Studies in Art Education. 12, (1), 11-16.

Gardner, H. (1971). The development of sensitivity to artistic styles. Journal of Aesthetics and Art Criticism, 29, 515 – 27.

Gardner, H. (1972a). Style sensitivity in children. Human Development. 15, 325 –338.

Gardner, H. (1972b). On figure and texture in aesthetic perception. The British Journal of Aesthetics. 12, (1), 40-59.

Gardner, H. (1973a). Children's sensitivity to musical styles. Merrill-Palmer Quarterly. 19, 67-77.

Gardner, H. (1973b). *The arts and human development*. New York: Wiley and Sons.

Gardner, H. (1979). *Development psychology after Piaget: an approach in terms of symbolization*. Human Development, 22, 73-88.

Gardner, H. (1993). *Frames of mind: the theory of multiple intelligences*. 2nd. edition. London: Fontana.

Gardner, H. (1994). *The arts and human development*. (paperback edition). New York: Basic Books.

Gardner, H; Phelps, E. & Wolf, D.P. (1990). The roots of adult creativity in children's symbolic products. in C.N. Alexander., & J. Langer, *Higher stages of human development*, Oxford: Oxford University Press.

Gardner, H.; Winner, E. & Kircher, M. (1975). *Children's conceptions of the arts*. The Journal of Aesthetic Education. 9, (3).

Garretson, R.L. (1966). *Music in childhood education*. New York: Appleton-Century-Crofts, Inc.

Geake, J.G (1999). An informational processing account of audiatonal abilities. Journal of Research Studies in Music Education. 12.

Gehring, A. (1910). *The basis of musical pleasure*. New York: Putnam.

Goodall. H. (2000). *Big bangs: the story of five discoveries that changed musical history*. London: Chatto and Windus.

Granier-Deferre, C.; Lecanuet, J-P.; Jacquet, A.Y. & Busnel, M.C. (1992). Prenatal discrimination of complex auditory stimulations. Poster presentation at the 8th. Meeting of the International Conference on Infant Studies, Miami.

Green, L. (1988). *Music on deaf ears*. Manchester: Manchester University Press.

Green, D.W. & Mitchell, D.C. (1978). The effects of context and content on immediate processing in reading. Quarterly Journal of Experimental Psychology. 30, 608-36.

Greer, R.D.; Dorow, L.G. & Hanser. S. (1973). Music discrimination training and the music selection behaviour of nursery and primary level children. Council for Research in Music Education, 35,(4) 30-43.

Greer, R.D.; Dorow, L.G. & Randall, A. (1974). Music listening preferences of elementary school children. Journal of Research in Music Education. 22. 284-291.

Greer, R.D.; Dorow, L.G.; Wachaus, G. & White, E.R. (1975). Adult approval and student's music selection behaviour. in C.K. Madsen; R.D.Greer., & C.H.Madsen, search *in music behaviour: Modifying music behaviour in the classroom*. New York: Teachers College Press.

Greer, R.D.; Randall, A. & Timberlake, C. (1971). The discriminate use of music listening as a contingency for improvement in vocal pitch acuity and attending behaviour. Council for Research in Music Education Bulletin. 26. 10-17.

Gregory, A.H. (1978). Perception of clicks in music. Perception Psychophysiology. 24. 171-4.

Gregory, A. H. (1997). The roles of music in society: the ethnomusicological perspective. in D.J. Hargreaves. & A.C. North. *The social psychology of music*. Oxford: Oxford University Press.

Gregory, D. (1993). Relationships between music training and music preferences. New Ways. 8(2), C-5.

Grossman, F. (1978). The effect of an examiner's reported academic achievement and/or physical condition on examiner's scoring of the WISC-R verbal IQ. Unpublished Ph.D Dissertation. University of Iowa.

Gudmundsdottir, H.R. (1999). Children's Auditory Discrimination of Simultaneous Melodies. Journal of Research in Music Education. 47,(2). 101-110.

Gurney, E. (1880). *The power of sound*. London: Smith, Elder.

Hair, H.I. (1981). Verbal identification of musical concepts. Journal of Research in Music Education. 29, 11-21.

Halpern, A. R. (1984). Organisation in memory for familiar songs. Journal of Experimental Psychology: Learning, Memory and Cognition. 10, 496-512.

Hanley, J.H. (1978). A comparative study of the sensitivity of the WISC and WISC-R to examiner and subject variables. Unpublished Ph.D dissertation. St.Louis University. Dissertation Abstract International 1978, 39,1480B-1481B. (University Microfilm order Number 7814571).

Hanslick, E. (1986). *On the musically beautiful: a contribution towards the revision of the aesthetic of music*. Indianapolis: Hackett Publishing Company.

Hargreaves, D.J. (1982). The development of aesthetic reactions to music. [Special Issue] Psychology of Music. 51-54.

Hargreaves, D.J. (1986). *The developmental psychology of music*. Cambridge: Cambridge University Press.

Hargreaves, D.J. (1986b). Developmental psychology and music education. Psychology of Music. 14, 83-96.

Hargreaves, D.J. (1999). Developing musical creativity in the social world. CRME Bulletin. 142, 22- 34.

Hargreaves, D.J.; Comber, C. & Colley, A. (1995). Effects of age, gender, and training on musical preferences of British secondary school students. Journal of Research in Music Education. 43, (3). 242-250.

Hargreaves, D.J. & North, A. (1997). *The social psychology of music*. Oxford: Oxford University Press.

Hargreaves, D.J. & North, A. (1999). *Developing concepts of musical style*. Musicae Scientiae. **2**. 193-213.

Hargreaves, D.J. & Zimmerman, M.P. (1992). Developmental theories of music learning. in R. Colwell, (ed.). *Handbook of research on music teaching and learning*. New York: Schirmer Books.

Harker, D. (1985). *Fakesong: The conspiracy of good taste*. Milton Keynes. Open University Press.

Harris, R. (1980). *The language makers*. New York: Ithaca.

Haslam, S.A.; Oakes, P.J.; McGarty, C.; Turner, J.C. & Onorato, R.S. (1995). Contextual Changes in the prototypicality of extreme and moderate outgroup members. European Journal of Social Psychology. **25**, 509-530.

Hata, Y., Tsudzuki, A., Kuze, T. & Emi, Y. (1958). Relationships between the tester and the subject as a factor influencing on the intelligence test score. Japanese Journal of Psychology. **29**, 95-99.

Hauser, A. (1962). *The social history of art*. (trans. in collaboration with the author by Stanley Godman) (1951), reprinted. London, (1962).

Heffernan, H. & Todd, V. (1960). *The kindergarten teacher*. Boston: D.C. Heath and Company.

Hekkert, P. & Snelders, D. (1996). Comment on Prototypicality Revisited. Empirical Studies of the Arts. **14**(1). 105-107.

Hekkert, P. & van Wieringen, P.C.W. (1990). Complexity and prototypicality as determinants of the appraisal of cubist paintings. British Journal of Psychology. **81**. 483-495.

Hewstone, M. & Manstead, A.S.R. (1995). Social Psychology in *The Blackwell encyclopaedia of social psychology*. (ed.) M. Hewstone and A.S.R. Manstead. Oxford: Blackwell.

Hewstone, M.; Stroebe, W. & Stephenson, G.M. (1996). *Introduction to Social Psychology*. 2nd. Edition. Oxford: Blackwell

Hill, M.M. (1966). "Hop-Skip-Jump into music appreciation", School and Community, **LIII** (December) **30**.

Hirst, P. H. (1972). Liberal education and the nature of knowledge. in R.D. Archambault. (Ed.) *Philosophical analysis and education*. London: Routledge and Kegan Paul.

Hornby, N. (1995). *High Fidelity*. New York: Riverhead Books.

Hutt, M.L.(1947). *A Clinical study of 'Consecutive' and 'adaptive' testing with the revised Stanford - Binet*. Journal of Consulting Psychology. **11**, 93-103.

Isenberg, S.J. & Bass, B.A. (1974). Effects of Verbal and Noverbal Reinforcement on the WAIS Performance of Normal Adults. Journal of Consulting and Clinical Psychology. **42**.(3), 467.

James, J. (1995). *The Music of the Spheres*. London: Abacus.

Karma, K. (1994) Auditory and visual temporal structuring: how important is sound to musical thinking. Psychology of Music and Music Education. **22**. (1).

Karpeles, M. (1973). *Cecil Sharp*. BBC Radio 4. January 7th.

Kastner, M.P. & Crowder, R.G. (1990). Perception of the major /minor distinction: emotional connotations in young children. Music Perception **8**, 189 –202.

Katz, R. (1964). *Body Language: a study in unintentional communication*. Unpublished Ph.D thesis. Harvard University.

Keddie. N. (1973). *Tinker, tailor: the myth of cultural deprivation*. Harmondsworth: Penguin.

Kellett, M. (2000). Raising musical esteem in the primary classroom: an exploratory study of young children's listening skills. British Journal of Music Education **17**, (2), 157-181.

Kennedy, W.A. & Vega, M. (1965). Negro's performance and examiners race. Journal of Personality and Social Psychology. **2**, (6). 839-843.

Keogh, B.K. & Macmillan, D.L. (1971). Effects of motivational and presentation conditions on digit recall of children of differing socio-economic, racial, and intelligence groups. American Education Research Journal **8**, 27-38.

Killian, J.N. (1990). Effect of model characteristics on musical preferences of junior high students. Journal of Research in Music Education **38**, 115-123.

Kindler, A.M. (1997). *Child development in art*. Reston, Va.: National Art Education Association.

Klugman, S.F. (1948). The effect of money incentive versus praise upon the reliability and obtained scores of the revised Stanford - Binet test. Journal of General Psychology. **30**, 255 – 269.

Konecni, V. J. (1979). Determinants of aesthetic preference and effects of exposure to aesthetic stimuli: Social, emotional and cognitive factors. in B.A. Maher (ed) *Progress in experimental personality research* (Vol.9). New York: Academic Press. 149-197.

Konecni, V. J. (1982). Social interaction and musical preference. in D. Deutsch. *The psychology of music*. New York: Academic Press.

Konecni, V. & Sargant - Pollock, D. (1976). Choice between melodies differing in complexity under divided attention conditions. *Journal of Experimental Psychology: Human Perception and Performance* **2**, 347 - 56.

Koopman, C. (1995). Stage theories of musical development. *Journal of Aesthetic Education* **29**,(2), 49-66.

Koroscik, J. (1997). Young people understanding works of art. in A.M. Kindler (ed.). *Child development in art*. Reston, Va.: National Art Education Association.

Kreitler, H. & Kreitler, S. (1972). *Psychology of the Arts*. Durham. N.C.: Duke University Press.

Kuhn, T.L. (1975). The effect of teacher approval and disapproval on attentiveness, musical achievement, and attitude of fifth grade students. in *Research In Music Behaviour: Modifying Music Behaviour in the Classroom*. 1975. Madsen, C.K., Greer, D.R., & Madsen, C.H. Jr. New York: Teachers College Press.

Labov, W. (1970). The logic of non-standard English. in F.Williams *Language and poverty*. Chicago: Markham.

Langer, S. (1953). *Feeling and Form*. New York: Scribener.

Langer, S. (1979). *Philosophy in a new key*. (3rd. edition). Cambridge: Harvard University Press.

Lantz, B. (1945). Some dynamic aspects of success and failure. *Psychological Monographs*. **59**.

Lawton, D. (1980). *The politics of the school curriculum*. London: Routledge and Kegan Paul.

Leach, R. & Palmer, R. (1978). *Folk music in schools*. Cambridge: Cambridge University Press.

Le Blanc. A. (1991). Effects of maturation/ aging on music listening preference: a review of the literature. Paper presented at the Ninth National Symposium on Research in music behaviour, Canon Beach. Oregon.

Le Blanc, A.; Sims, W.L.; Siivola, C. & Obert, M. (1993). Musical style preferences of different - age listeners. Paper presented at the Tenth National Symposium on Research in Music Behaviour, University of Alabama, Tuscaloosa, Alabama. U.S.A

Lecanuet, J-P. (1996). Prenatal auditory experience. in I. Deliege and J. Sloboda (eds.). *Musical Beginnings: Origins and development of musical competence*. Oxford: Oxford University Press.

Leppert, R. (1987). Music, domestic life and cultural chauvinism: images of British subjects at home in India, in R. Leppert and S. McClary (eds.). (1987). *Music and society: The politics of composition, performance and reception*. Cambridge: Cambridge University Press.

Leppert, R., & McClary, S. (eds.). (1987). *Music and society: The politics of composition, performance and reception*. Cambridge: Cambridge University Press.

Lerdahl, F. & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, Ma.: MIT Press.

Light, P.; Littleton, K.; Messer, D. & Joiner, R. (1994) Social and communicative processes in computer-based problem solving. European Journal of Psychology of Education, 9, 93-109.

Linblom B. & Sundberg, J. (1970). Towards a generative theory of melody. Swedish Journal of Music, 52, 71-88.

Lindner, G.A. (1871). *Ideen zur psychologie der gesellschaft als grundlage der sozialwissenschaft*. Vienna: Gerold.

Livesley, W.J. & Bromley, D.B. (1973). *Person perception in childhood and adolescence*. London: Wiley.

Locke, S. & Keller, L. (1973). Categorical perception in a non-linguistic mode. Cortex, 9, 355-69.

Longuet-Higgins, H.C. (1976). The perception of melodies. Nature, 263, 646-653.

Lundin, R.W. (1967) *An objective psychology of music*. (2nd edition.). New York: Ronald Press.

Lutey, C. & Copeland, E. (1982). Cognitive assessment of the school age child in C. Reynolds., & T. Gutkin. (eds.). *The handbook of school psychology*. 2nd. Edition. New York: Wiley.

Machotka, P. (1966). Aesthetic criteria in childhood: justifications of preference. Child Development, 37, 877-85.

Madsen, C.K.; Wolfe, D.E. & Madsen, C.H. (1969). The effect of reinforcement and directional scalar methodology on intonational improvement. Journal of the Council for Research in Music Education. **18**, 22-34

Madsen, C.K.; Wolfe, D.E. & Madsen, C.H.Jr. (eds.). (1975). *Research in Music Behaviour*. New York: Teachers College Press.

Mandler, G., & Sarason, S.B. (1952). A study of anxiety and learning. Journal of Abnormal and Social Psychology. **47**, 166-173.

Martin, P. J. (1995). *Sounds and society: themes in the sociology of music*. Manchester: Manchester University Press.

Martindale, C. (1990). *The clockwork muse: the predictability of artistic change*. New York: Basic Books.

Martindale, C. (1996). A note on the relationship between prototypicality and preference. Empirical Studies of the Arts. **14**,(1). 109-113.

Martindale, C., Moore, K. & West, A. (1988). Relationship of preference judgements to typicality, novelty, and mere exposure. Empirical Studies of the Arts **6**, 431-55.

Masling, J. (1965). Differential indoctrination of examiners and Rorschach responses. Journal of Consulting Psychology **29**, 198-201.

May, W.V. (1985). Musical style preferences and aural discrimination skills of primary grade school children. Journal of Research in Music Education, **33**, 7-22.

McCarthy, D. (1944). A study of the reliability of the Goodenough drawing test of intelligence. Journal of Psychology **47**, 166-173.

McDougall, W. (1908). *Introduction to social psychology*. London: Methuen.

McPherson, G.E. (1995). The assessment of musical performance: development and validation of five new measures. Psychology of Music and Music Education. **23**,(2) 142-159.

Mead, G.H. (1934). *Mind, self and society*. Chicago: University of Chicago Press.

Messenger, J. (1958). Esthetic talent. Basic College Quarterly **4**, 20-4.

Meyer, L. B.(1956). *Emotion and, meaning in music*. Chicago: Chicago University Press.

Meyer, L.B.(1989). *Style and music: theory, history, and ideology*. Philadelphia: University of Pennsylvania Press.

Middleton, R. (1990). *Studying popular music*. Milton Keynes: Open University Press.

Milliman, R.E. (1982). Using background music to affect the behaviour of supermarket shoppers. Journal of Marketing 46, 86-91.

Mills, J. (1998). *Music in the Primary School*. (revised edition). Cambridge: Cambridge University Press.

Mintz, N. (1957). On the psychology of aesthetics and architecture. Unpublished paper. Barndeis University.

Moog, H. (1976). The musical experience of the pre-school child. (trans. C. Clarke). London: Schott.

Moon, C.; Cooper, R.P., & Fifer, W.P. (1993). Two-day olds prefer their native language. Infant Behaviour and Development 16, 495-500.

Moore, C.L. & Retish, P.M. (1974). Effect of the examiners' race on black children's Wechsler preschool and primary scale of intelligence IQ. Developmental Psychology 10, 672-676.

Morales, E.S. (1977). Examiner effects on the testing of Mexican-American bilingual children in the early elementary grades. Unpublished Ph.D dissertation. Texas Tech University.

Morrongiello, B.A. (1986). Infants' perception of multi-group auditory patterns. Infant Behaviour and Development 9, 307-319.

Morrongiello, B.A. (1992). Effects of Training on Children's Perception of Music: A Review. Psychology of Music. 20, 29-41.

Morrongiello, B.A.; Roes, C.L. & Donnelly, F. (1989). Children's perception of musical patterns: effects of music instruction. Music Perception. 6, 447-462.

Mugny, G.; Levy, M. & Doise, W. (1978). Conflit sociocognitif et developpement cognitif. Revue Suisse de Psychologie Pure et Appliquee 37, 22-43.

Murdock, G. & Phelps, G. (1972). Youth Culture and the school revisited. British Journal of Sociology. 23, 478-482.

Murray, K.C. (1975). The effect of teacher approval / disapproval on the performance level, attentiveness, and attitude of high school choruses. In C.K. Madsen, R.D. Greer, & C.H. Madsen, Jr. (eds.). *Research in music behaviour: modifying music behaviour in the classroom*. New York: Teachers College Press.

Nattiez, J-J. (1990). *Music and discourse: towards a semiology of music*. Princeton; New Jersey: Princeton University Press.

Newman, T. (1966). Music for the Pre-school child. American Music Teacher. XV, (January) 15.

Nietzsche, F. (1967). *The birth of tragedy and The case of Wagner* trans. W. Kaufmann. New York: Vintage Books

North, A.C. & Hargreaves, D.J. (1996). Responses to music in a dining area. Journal of Applied Social Psychology. 26, 491 – 501.

North, A.C. & Hargreaves, D.J. (1999). Music and driving game performance. Scandinavian Journal of Psychology 40, 285-292.

North, A.C. & Hargreaves, D.J. (2000). Musical preferences during and after relaxing and exercising. American Journal of Psychology 113, 43-67.

North, A.C.; Hargreaves, D.J. & Heath, S.J. (1998). Musical tempo and time perception in a gymnasium. Psychology of Music 26, (1).

O'Hare, D. (ed.). (1981). *Psychology and the arts*. Brighton: Harvester Press.

Olsho, L.W.; Koch, E.G.; Halpin, C.F. & Carter, E.A. (1987). An observer-based psychoacoustical procedure for use with young infants. Developmental Psychology. 23, 627-40.

O'Neill, S.A. & Sloboda, J.A. (1997). The effects of failure on children's ability to perform a musical test. Psychology of Music. 25, 18-34.

Papousek, M. & Papousek, H. (1981). Musical elements in the infant's vocalizations: their significance for communication, cognition and creativity in Advances in Infancy Research 1, (ed. L.P. Lipsitt), 163-224. Norwood, New Jersey: Ablex.

Papousek, H. (1996a). Musicality in infancy research. in I. Deliège and J. Sloboda (eds.). (1996). *Musical beginnings: origins and development of musical competence*. Oxford: Oxford University Press.

Papousek, M. (1996b). Intuitive parenting: a hidden source of musical stimulation in infancy. in I. Deliège and J. Sloboda (eds.). (1996). *Musical beginnings: origins and development of musical competence*. Oxford: Oxford University Press.

Parsons. M.J.(1987). *How we understand art: a cognitive developmental account of aesthetic experience*. London: Cambridge University Press.

Partch, H.(1974). *Genesis of a music*. (2nd edition) New York: Da Capo Press.

Pedersen, D.M.; Shinedling, M.M. & Johnson, D.L. Effects of sex of examiner and subject on children's quantitative test performance. Journal of Personality and Social Psychology. **10**, (3). 251-254.

Persson, R.S. & Robson, C. (1995). The limits of experimentation: on researching music and musical settings. Psychology of Music. **23**, 39-47.

Pflederer, M. (1964). The responses of children to musical tasks embodying Piaget's principle of conservation. Journal of Research in Music Education. **12**, 251-268.

Pflederer Zimmerman, M. (1967). Conservation and the development of musical intelligence. Journal of Research in Music Education. **15**, 215-23.

Pflederer Zimmerman, M. & Sechrest, L. (1970). Brief focused instruction and musical concepts. Journal of Research in Music Education **18**, 25-36.

Phillips, J. (1966). Performance of father-present and father-absent Southern Negro boys on a simple operant task as a function of the race and sex of the experimenter and the type of social reinforcement. Doctoral Dissertation, University of Minnesota.

Plummeridge, C. (1991). *Music education in theory and practice*. London: Falmer Press.

Pole, W. (1877). Series of lectures published in Fraser's Magazine.

Pole, W. (1924). *The philosophy of music*. London: Kegan Paul, Trench, Trubner.

Pollard, A. (1996). *The social world of children's learning*. London: Cassell.

Pollard-Gott, L. (1983). Emergence of thematic concepts in repeated listening to music. Cognitive Psychology **15**, 66-94.

Quay, L.C. (1971). Language dialect, reinforcement, and the intelligence-test performance of Negro children. Child Development **42**, 5-15.

Quereshi, M.Y. (1968). Intelligence test scores as a function of sex of experimenter and sex of subject. Journal of Psychology **69**, 277-284.

Radocy, R.E. (1975). A naïve minority of one and deliberate majority mismatches of tonal stimuli. Journal of Research in Music Education **23**, 120-33.

Radocy, R.E. (1976). Effects of authority figure biases on changing judgements of musical events. Journal of Research in Music Education. **24**, (3). 119-128.

Radocy, R.E. & Boyle, J.D. (1997). *Psychological foundations of musical behaviour*. (3rd. Edition). Springfield. Illinois: Charles C. Thomas.

Read, H. (1955). *Icon and idea*. London: Faber and Faber.

Reichenberg-Hackett, W. (1953). Changes in Goodenough drawings after a gratifying experience. American Journal of Orthopsychiatry **23**, 501-517.

Reid, L.A. (1961). *Ways of knowledge and experience*. London: Allen and Unwin.

Reimer, B. (1989). *A philosophy of music education*. (2nd. Edition). Englewood Cliffs. N.J.: Prentice Hall.

Repp, B. (1984). Categorical perception: issues, methods, findings. in N.J. Lass. (ed.). *Speech and language: advances in basic research and practice*. (Vol.10). New York: Academic Press.

Revesz, G. (1953). *Introduction to the psychology of music*. London: Longmans.

Riecken, H.W. (1962). A program of research on experiments in social psychology. in N.F.Washburne. (ed.). *Decisions, values and groups*. Vol. 2. New York: Pergamon Press. 25-41.

Rigg, M.C. (1948). Favourable versus unfavourable propaganda in the enjoyment of music. Journal of Experimental Psychology. **38**, 78-81.

Risset, J-C. & Wessel, D.L. (1982). Exploration of timbre by analysis and synthesis. in D.Deutsch, (ed.). (1982). *The psychology of music*. New York: Academic Press.

Rodriguez, C.X. & Webster, P.R. (1997). Development of children's verbal interpretative responses to music listening. Bulletin of the Council for Research in Music Education. **134**, 9-30.

Rogers, V.R. (1957). Children's musical preferences as related to grade level and other factors. Elementary School Journal **57**, 433-5.

Rosenthal, R. (1964). The effect of the experimenter on the results of psychological research in B.A. Maher (ed.). *Progress in experimental personality research*. Volume 1. New York: Academic Press. 79-114.

Rosenthal, R. (1966). *Experimenter effects in behavioural research*. New York: Appleton.

Rosenthal, R. (1976). *Experimenter effects in behavioural research*. (Enlarged edition). New York: Halstead Press.

Rosenthal, R. & Halas, E.S. (1962). Experimenter effect in the study of invertebrate behaviour. Psychological Reports. **11**, 251-256.

Rosenthal, R., & Jacobson, L. (1966). Teachers' expectancies: determinants of pupils' IQ gains. Psychological Reports. 19, 115-118.

Rosenthal, R. & Jacobson, L. (1968). *Pygmalion in the classroom: teacher expectations and pupils' intellectual development*. New York: Holt, Rinehart and Winston, Inc.

Ross, E.A. (1908). *Social psychology*. New York: Macmillan.

Sachs, C. (1943). *The rise of music in the ancient world*. New York: Norton.

Salzer, F. (1952). *Structural hearing*. New York: Charles Boni.

Samuel, W.; Soto, D.; Park, M.; Ngissah, P. & Jones, B. (1976). Motivation, race, social class and I.Q. Journal of Educational Psychology 68, 273-285.

Samuel, W. (1977). Observed I.Q. as a factor of test atmosphere, tester expectation and race of tester. A replication of female subjects. Journal of Educational Psychology. 77, (30). 207-212.

Sarason, I.G. & Buss, A.H. (1965). Test anxiety and experimental conditions. Journal of Social Psychology. 1, 499-505.

Sarason, I.G. & Harnatz, M.G. (1965). Test anxiety and experimenter conditions. Journal of Personality and Social Psychology. 1, 499-505.

Sattler, J.M. (1970). Racial 'experimenter effects' in experimentation, testing, interviewing, and psychotherapy. Psychological Bulletin 73, (2) 137-160.

Sattler, J.M. & Ryan, J.J. (1973). Scoring agreement on the Stanford-Binet. Journal of Clinical Psychology. 29, 35-38.

Sattler, J.M. & Theye, F. (1967). Procedural, situational, and interpersonal variables in individual intelligence testing. Psychological Bulletin. 68, (5). 347-360.

Schaie, K.W. (1965). A general model for the study of developmental problems. Psychological Bulletin 64, 92-107.

Schroeder, H.E. & Kleinsasser, L.D. (1972). Examiner bias: a determinant of childrens' verbal behaviour on the WISC. Journal of Consulting Psychology. 39, 451-454.

Schuckert, R.F. & McDonald, R.L. (1968). An attempt to modify the musical preferences of pre-school children. Journal of Research in Music Education. 16, 39-45.

Schuessler, K.F. (1948). Social background and musical taste. American Sociological Review 13, 330-5.

Schwarz, M.L. (1966). The scoring of WAIS Comprehension responses by experienced and inexperienced judges. Journal of Clinical Psychology 22, 425-427.

- Seashore, C.E. (1919). *Manual of instructions and interpretations of measures of musical talent*. Chicago: C.H. Stoelting.
- Seashore, C.E.(1938). *Psychology of music*. New York: Dover Publications.
- Serafine. M.L. (1980)*Music as cognition: the development of thought in sound*. New York: Columbia University Press
- Sharp, C. J. (1907). *English folk song: some conclusions*. London: Novello.
- Sharp, C. J.(1912). *Folk singing in schools*. London: English Folk Dance and Song Society.
- Shepard, R.N. (1982). Structural representations of musical pitch. in D.Deutsch, (ed). (1982). *The psychology of music*. New York: Academic Press.
- Shepherd, J. (1977). The 'meaning' of music, in J. Shepherd, P. Virden, G. Vulliamy, & T. Wishart. (eds.). *Whose Music? A sociology of musical languages*, London: Latimer
- Shuter - Dyson, R. & Gabriel, C. (1981). *The psychology of musical ability*. London: Methuen.
- Sibley, F. (1965). Aesthetics and nonaesthetics. Philosophical Review. 74.
- Sibley, F. (1959). Aesthetic concepts. Philosophical Review. 68.
- Simmel, G. (1908). *Soziologie – untersuchungen uber die formen der vergesellschaftung*. Leipzig: Duncker & Humblot.
- Simon, H.A. & Sumner, R.K. (1968). Pattern in Music. in B.Kleinmuntz (ed.), *Formal representation of human judgement*. New York: John Wiley.
- Sloboda, J. (1974). The eye-hand span: An approach to the study of sight reading. Psychology of Music 2, 4-10.
- Sloboda, J. (1977). Phrase units as determinants of visual processing in music reading. British Journal of Psychology 68, 117-24.
- Sloboda, J. (1982). Music performance. in D. Deutsch (ed.). *The psychology of music*. New York: Academic Press.
- Sloboda, J. (1985). *The musical mind*. Oxford: Clarendon Press.
- Sloboda, J. (1988). *Generative processes in music*. Oxford: Clarendon Press.

Sloboda, J. (1989). Music as a language. in F. Wilson, & F. Roehmann (eds.). *Music and child development*. St.Louis: MMB.

Sloboda, J. (1991). Music structure and emotional response: Some empirical findings. *Psychology of Music*. **19**, (2). 110-120.

Sloboda, J. (1994). *Music performance* in R. Aiello and J. Sloboda, *Musical perceptions*. Oxford: Oxford University Press.

Sloboda, J. & Davidson, J. (1996). The young performing musician. in I. Deliege, and J. Sloboda, (eds.). *Musical beginnings. origins and development of musical competence*. Oxford: Oxford University Press.

Sluckin, W.; Hargreaves, D.J. & Colman, A. M. (1982). Some experimental studies of familiarity and liking . *Bulletin of the British Psychological Society*. **35**, 189-94.

Small. C. (1977). *Music - Society - Education*. London: John Calder Ltd.

SPSS. (1998). *SPSS for Windows*. New Jersey: Prentice Hall.

Stevenson, H.W. (1961). Social reinforcement with children as a function of Child Age, sex of experimenter and sex of subject. *Journal of Abnormal Social Psychology*. **63**, 147-154.

Stevenson, H.W. & Allen, S. (1964). Adult performance as a function of sex of experimenter and sex of subject. *Journal of Abnormal Social Psychology*. **68**, 214-216.

Stevenson, H.W.; Keen, R. & Knight, R.M. (1963). Parents and strangers as reinforcing agents for childrens' performance. *Journal of Abnormal Social Psychology*. **67**, 183-186.

Stevenson, H.W. & Odom, R.D. (1963). Visual reinforcement with children. Unpublished Manuscript. University of Minnesota.

Storr, A. (1960). *The integrity of the personality*. Harmondsworth: Penguin Books.

Storr, A. (1993). *Music and the mind*. London: Harper Collins.

Summers, G.F. & Hammond, A.D. (1965). Toward a paradigm of respondent bias in survey research. Unpublished paper. University of Wisconsin.

Sunberg, J. & Lindblom, B. (1976). *Generative theories in language and music descriptions*. *Cognition*. **4**, 99-122.

Swanson, B.R. (1962). *Music in the education of children*. Belmont, California: Wadsworth Publishing Company Inc.

- Swanwick, K. (1968). *Popular music and the teacher*. Oxford: Pergamon Press.
- Swanwick, K. (1979). *A basis for music education*. London: Routledge.
- Swanwick, K. (1988). *Music, mind, and education*. London: Routledge.
- Swanwick, K. (1994). *Musical knowledge: intuition, analysis and music education*. London: Routledge.
- Swanwick, K. & Tillman, J. (1986). The sequence of musical development. British Journal of Music Education. **3**, (3), 305-39.
- Szczelkun, S. (1993). *The conspiracy of good taste: William Morris and Cecil Sharp*. London: Woking Press.
- Tafuri, J.; Addressi, A.; Luzzi, C. & Baroni, M. (1994). *The development of musical stylistic competence in children*. Paper to 15th. International Research Seminar.
- Tarrant, M.; North, A.C. & Hargreaves, D. (2000). English and American adolescents' reasons for listening to music. Psychology of Music. **28**, (2)
- Taubman, H. (1958). *How to bring up your child to enjoy music*. Garden City, New York: Hanover House.
- Terwogt, M.M. & Van Grinsven, F. (1991). Musical expression and mood states. Psychology of Music. **19**, 99-109.
- Teplov, B.M. (1966). *Psychologie des aptitudes musicales*. Paris: Presses Universitaires de France.
- Thorisson, T. (1998). Comparison of novice listeners' similarity judgements and style categorisation of classic and romantic piano exemplars. Psychology of Music. **26**, (2) 186-196.
- Thorpe, L.; Trehub, S.; Morrongiello, B.A. & Bull, D. (1988). Perceptual grouping by infants and pre-school children. Developmental Psychology. **24**, 484-491.
- Tiber, N. & Kennedy, W.A. (1964). The effects of incentives on the intelligence test performance of different social groups. Journal of Consulting Psychology. **28**, 187.
- Tolbert, E. (2001). Music and meaning: an evolutionary story. Psychology of Music. **29**, (1), 84-94.
- Trehub, S.E. (1985). Auditory pattern perception in infancy. in S.E. Trehub and B.A. Schneider (Eds.) *Auditory development in infancy*. (pp.183-195). New York: Plenum.

Trehub, S.E. (1987). Infants' perception of musical patterns. Perception and Psychophysics. **41**, 635-641.

Trehub, S.; Cohen, A.; Thorpe, L. & Morrongiello, B.A. (1986). Development of the perception of musical relationships: Semitone and diatonic structure. Journal of Experimental Psychology: Human Perception and Performance **12**, 295-301.

Trehub, S. & Thorpe, L. (1989). Infants' perception of rhythm: Categorisation of auditory sequences by temporal structure. Canadian Journal of Psychology. **43**, 217-229.

Trehub, S.E.; Thorpe, L.A. & Morrongiello, B.A. (1985). Infants' perception of melodies: changes in a single tone. Infant Behaviour and Development. **8**, 213-223.

Upitis, R. (1987). *Children's understanding of rhythm: The relationship between development and music training*. Psychomusicology. **7**, 41-60.

Vulliamy, G. (1977). Music and the mass culture debate, in J. Shepherd, P. Virden, G. Vulliamy and T. Wishart. *Whose Music? A sociology of musical languages*. London: Latimer.

Vulliamy, G. & Shepherd, J. (1984). Sociology and music education: a response to Swanwick, British Journal of the Sociology of Education **5**.

Vygotsky, L. S. (1978). *Mind in society: the development of higher psychological processes*. Cambridge, Mass.: Harvard University Press.

Vygotsky, L. S. (1997). *Thought and language*. Cambridge, Mass.: The MIT Press.

Wadsworth, B.J. (1996). *Piaget's theory of cognitive and affective development: foundations of constructivism*. 5th. Edition. London: Longmans.

Walker, A. (1983). The contribution of Walter Carrol (1869 –1955) to music education. Unpublished M.Ed Thesis. University of Manchester.

Walker, A.(1989). *Walter Carroll: The children's composer*. Manchester: Forsyth.

Wang, C. & Salzburg, R. (1984). Discrimination of modulated music tempo by string students. Journal of Research in Music Education. **32**, 123-131.

Ward, D. (1984). Personal construct theory: its application to research in music education and therapy. In M.Ross (ed.). Artstrip, **14**. Exeter: University of Exeter Publications. 197-209.

Weiss, N.A. (1995). *Introductory statistics*. (4th. Edition). New York: Addison-Wesley.

Wertsch, J.V. (1985). *Vygotsky and the social formation of mind*. Cambridge, Mass.: Havard University Press.

Whitfield, W.A. & Slatter, P.E. (1979). The effects of categorization and prototypicality on aesthetic choice in a furniture selection task. British Journal of Psychology 70, 65-75.

White, J. (1997). *Do Howard Gardner's multiple intelligences add up?* London: Bedford Way Papers, Institute of Education.

Williams, R. (1972). Effects of musical aptitude, instruction, and social status on attitudes toward music. Journal of Research in Music Education. 20, (3).

Wing, H.D. (1968). *Tests of musical ability and appreciation*. (2nd edition). British Journal of Psychology, Monograph Supplement. 27, Cambridge: Cambridge University Press.

Winner, E. (1982). *Invented worlds: the psychology of the arts*. Cambridge, Mass.: Havard University Press.

Winneykamen, F. (1990). *Appendre en imitant?* Paris: Presses Universitaire de France.

Wishart, T. (1977). Musical writing, musical speaking, in J. Shepherd, P. Virden, G. Vulliamy and T. Wishart. (eds.). *Whose Music? A sociology of musical languages* London: Latimer Press.

Wolff, J. (1987). The ideology of autonomous art. in R. Leppert and S. McClary. *Music and Society*. Cambridge: Cambridge University Press.

Wolpert, R. (1990). Recognition of melody, harmonic accompaniment, and instrumentation: Musicians vs. non musicians. Music Perception 8, 95-105.

Wood, A. & Coltman, P. (1973). Talking mathematics. In E. Bearne (ed.). *Use of language across the primary curriculum*. London: Routledge.

